

타원형 정보와 웨이블릿 패킷 분석을  
이용한 얼굴 검출 및 인식

연세대학교 대학원  
생체공학협동과정  
전기전자공학전공  
정 명 호

타원형 정보와 웨이블릿 패킷분석을  
이용한 얼굴 검출 및 인식

지도 박 민 용 교수

이 논문을 석사 학위논문으로 제출함

2003년 6월 일

연세대학교 대학원

생체공학협동과정

전기전자공학전공

정 명 호

# 정명호의 석사 학위논문을 인준함

심사위원 \_\_\_\_\_ 인

심사위원 \_\_\_\_\_ 인

심사위원 \_\_\_\_\_ 인

연세대학교 대학원

2003년 6월 일

# 감사의 글

---

새로운 마음가짐으로 시작한 대학원 생활. 지하철 운행이 힘들 정도로 유난히 비가 많이 오는 가운데 연구실 행사에 처음 참석한 2001년 7월 여름 콜로키움의 기억이 생생한 데 어느덧 2년이란 시간이 지났네요. 우선 주위에서 저를 지켜봐 주신 모든 분들에게 감사의 말을 전합니다.

학문의 길에 접어들 수 있는 길을 열어주시고 음으로 양으로 항상 따뜻한 관심으로 보살펴 주신 박민용 교수님께 감사 드립니다. 그리고 부족한 논문을 위해 많은 조언과 지도를 해주신 이철희 교수님과 김은태 교수님께 감사 드립니다. 2년이란 시간동안 학업을 지도해 주신 이명호 교수님, 박진배 교수님, 손광훈 교수님, 차일환 교수님, 박상희 교수님, 유선국 교수님, 서활교수님 및 여러 교수님들께도 고마운 마음을 전합니다.

연구실을 졸업하신후에도 항상후배들을 격려하고 도와 주시고 연구실 행사에 참석 하시는 박민기 선배님, 이창훈 선배님, 가민호 선배님, 노홍식 선배님, 조영완 선배님, 이기철 선배님, 권철 선배님, 이영희 선배님 및 모든 선배님들께 이 글을 빌려 감사한 마음을 전하고 싶습니다. 그리고 같이 생활하다 졸업한 민식형, 은호형, 한진형, 졸업생중에 가장 오래 같이 지낸 졸업생들. 후배들을 도와주고자 노력하신 성렬형, 어엿한 박사 이빠 창우와 도움만 받은 명환 그리고 짝궁이라 불릴 정도로 붙어다니 까만 해군 점효형, 민화를 유난히 좋아했던 삼성맨 창우형 그리고 도시같은 영주와 파워풀한 정기형도 생각이 나네요. 힘든 결정이었지만 회사생활을 그만두고 학업의 길을 갈수 있도록 도움을 준 경현형, 대희형, 김문겸과장님에게도 감사한 마음을 전합니다.

어렵고 힘들때마다 버팀목이 되어준 너따모 사람들. 너따모의 만형 철홍형, 항상 웃음과 힘을 준 준상형, 어렵고 힘들때마다 조언을 아끼지 않은 상일형, 지금은 약해졌지만 유머러스했던 철진형, 크리스마를 가진 불의를 참치 못하는 대영형, 웃음소리가 나보다도 멋지고(?) 얼큰이의 대부 윤형, 이가 이빠 극선형, 봉석, 학부 부터 같이 지낸 멋진 상수와 이벤트 매니저이자 자신보다 항상 남을 먼저 배려하는 상현, 매사에 최선을 다하는 태희, 어디로 튈지 모를 것 같지만 너따모 회원중 근대 계급이 가장 높은 민호, 울산을 지키고 있는 승명, 그리고 여러 후배 재완, 흥민, 점수, 태진, 정준에게도 저의 감사의 글을 올립니다.

그리고 친구로서 오랜시간 같이 지내오면서 옆에서 지켜봐준 원석, 영섭이의 우정도 고맙다라는 짧은 글로 대신하면서 멋진 친구들의 변치않은 우정과 무궁한 발전이 있기를 기도 드립니다.

2년여의 시간동안 같은 연구실에서 지내면서 졸업을 같이 하게되는 현석형도 축하드리고, 매사에 최선을 철완형, 고등학교 선생님인 석윤형, 문무를 겸비한 신형형, 그리고 언제나 말이 없지만 누구보다 부지런한 상형형, 때론 표정하나 말 한마디로 모든 연구실원을 웃음의 도가니로 만드는 재훈형, 그리고 공군 소령이자 모든일을 조용히 묵묵히 일하시면서 고생하시는 경중형과 운동이면 운동, 공부면 공부, 민승인 활력소 창호, 논문 준비하느라 바쁘지만 연구실의 대소사를 일일이 챙기는 꽃 혜령, 연구실의 키사노비이자 멋스러운 성완, 목소리가 유난히 크면서 통도 큰 친구 용원이, KIST의 언제나 활기찬 '섭섭 형제' 우섭, 범섭, 어딘지 모르게 우수한 청국장 같은 사나이 종근이와 '느끼남' 이라 불리우지만 마스크트 같은 일원이, 작고 탄탄한 사나이 유근. 그리고 새롭게 시작을 준비하는 경민이와 IT를 짊어지고 갈 재필이와 희성이 그리고 굶은일의 하느라 힘든 나현씨에게도 감사의 말을 전하며 앞으로 행복한 연구실 생활이 되기를 바랍니다.

그리고 언제나 따뜻한 눈길로 대해 주시면서 학업에 매진할수 있도록 도와 주신 장모님과 상준이 그리고 삼촌,숙모께도 감사 드리며, 나의 하나 밖에 없는 동생 명희와 매제 그리고 깜직한 귀염둥이 미선이 한테도 글로나마 모든 감사한 마음을 이글로 대신합니다.

어렸을때부터 이찌 주신 작은 아버지와 어머니 그리고 7월에 군대에 가는 상철이와 곧 대학생이 될 명욱이에게도 감사의 글을 올립니다.

그리고 마지막으로 30여년 이란 시간동안 못난 아들을 믿고, 물심양면으로 희생만을 해 오시며, 뒷바라지 해주시고 부족한 저를 항상 변함없는 신뢰와 가르침으로 사랑해주신 아버님과 밤낮 걱정으로 보이지 않는곳에서 사랑을 아끼시지 않는 어머님과 언제나 나의 편에서 영원한 동반자이자 나의 투쟁을 받아 주면서 불평없이 나를 이찌 주고 위로해 주면서 지켜봐준 사랑스러운 선영이와 앞으로의 세상의 밝은 빛을 보게될 우리 아가와 저의 작은 결실을 함께 하고자 합니다.

2003년 여름

정명호 드림

# 차 례

---

그 립 차 례 .....	iii
표 차 례 .....	iv
국 문 요 약 .....	v
제 1 장 서 론 .....	1
제 2 장 기존의 얼굴 검출 및 인식에 방법 .....	5
제 2.1 절 얼굴 검출 및 인식 방법들 .....	5
2.1.1 특징 기반 연구 .....	8
2.1.2 영상 기반 방법 .....	10
제 3 장 얼굴 검출 및 인식 기법 .....	15
제 3.1 절 타원정보로 얼굴영역 검출기법. ....	15
3.1.1 에지 검출과 에지 연결과정. ....	15
3.1.2 타원형 찾기 .....	17
제 3.2 절 PCA를 이용한 얼굴 인식 방법. ....	19
제 3.3 절 Wavelet 이론 .....	24
제 4 장 제안된 시스템의 구현. ....	28
제 4.1 절 얼굴 검출 알고리즘 .....	29
4.1.1 전처리 .....	30
4.1.2 얼굴 영상의 웨이블릿 변환 .....	32
4.1.3 형판정합을 이용한 얼굴 검출 .....	34
4.1.4 눈 영역 검출 알고리즘 .....	35

4.1.5 얼굴 영역 검출 .....	36
제 4.2 절 얼굴 인식 알고리즘 .....	37
4.2.1 웨이블릿 패킷 분해( Wavelet Packet Decomposition ) .....	37
4.2.2 얼굴영상에서의 웨이블릿 패킷분해 .....	38
4.2.3 얼굴 특징 국부화. ....	41
4.2.4 특징 벡터 추출 .....	44
제 5 장 실험 및 결과. ....	46
제 5.1 절 실험 환경 및 데이터 베이스 .....	46
제 5.2 절 개인 얼굴 식별 실험 .....	49
제 5.3 절 얼굴 표정 변화시 인식을 실험 .....	50
제 5.4 절 PCA 방법과의 실험 비교 .....	51
제 6 장 결론 및 추후 과제 .....	53
참 고 문 헌 .....	54
영 문 초 록 .....	59

## 그림 차례

---

그림 2-1	얼굴인식을 위해서 훈련에 사용되는 전형적인 이미지들	5
그림 2-2	실제적으로 적용할 얼굴 영상	6
그림 2-3	방법들에 따라서 나누어진 얼굴검출 및 인식방법	6
그림 2-4	ORL 데이터 베이스의 고유얼굴 예	11
그림 2-5	Sung 및 Poggio의 시스템에서의 분류	12
그림 2-6	Rowley et al.의 시스템	13
그림 2-7	Rowley et al.과 Sung 및 Poggio에 의해 적용된 전처리 방법	13
그림 3-1	주성분 분석을 위해 훈련 집합안에 영상들	20
그림 3-2	평균 얼굴	20
그림 3-3	원영상에서 평균을 뺀 영상	21
그림 3-4	고유얼굴들	22
그림 3-5	입력영상	23
그림 3-6	인식된 얼굴	23
그림 4-1	얼굴 검출 및 인식 알고리즘	28
그림 4-2	얼굴 검출 알고리즘.	30
그림 4-3	제안된 밝기 조정후의 영상들	32
그림 4-4	얼굴 형판 생성과정	34
그림 4-6	눈 영역 추출	36
그림 4-7	웨이블릿과 Frakas식을 적용한후 얼굴 영역의 검출	37
그림 4-8	웨이블릿에 적용할 H,G 필터 H(실선) G(점선)	39
그림 4-9	웨이블릿 패킷 가지	40
그림 4-10	웨이블릿 패킷 가지의 2단계를 통해 얻어진 부영상들	40
그림 4-11	FERET 데이터로부터 추출된 얼굴들	42
그림 4-12	저주파 얼굴의 분리후 영상	43



그림 5-1 MIT FACES DATA의 예 .....	47
그림 5-2 FERET DATA의 예 .....	48
그림 5-3 FACES와 FERET 영상의 여섯개의 고유얼굴들 .....	51

## 표 차 례

---

표 4-1 조명에 따른 영상 상태 .....	31
표 5-1 실험 데이터와 실험 환경 .....	46
표 5-2 FACES 영상 수에 따른 인식률 .....	49
표 5-3 FERET 영상에 따른 인식률 .....	49
표 5-4 MIT 데이터 베이스의 얼굴 표정 변화 입력시 인식율 .....	50
표 5-5 PCA방법과 제안한 방법과의 비교 .....	52

## 국 문 요 약

---

### 타원형 정보와 웨이블릿 패킷 분석을 이용한 얼굴 검출 및 인식

최근 들어 얼굴 검출 및 인식 기술은 개인의 정보 보호 및 신원 확인에 의해 발전하고 있다. 얼굴 검출의 방법은 신체적인 접촉이 적고, 사람 얼굴의 특이성 때문에 다양한 방법으로 오류율이 적고 검출 및 인식이 정확하도록 최근 10여년간 발전하고 있다.

얼굴 검출 및 인식 분야는 영상신호처리, 영상 분할, 신경회로망 또는 통계적 패턴인식 기술등을 종합하는 기술로서 선진 외국의 경우, 기업은 물론 정부 심지어 국제 협력기구의 주도하에 1970년대부터 학제적 공동연구가 본격 전개되기 시작하였으며 현대 유럽의 경우 여러 국가에서 공동으로 연구 단체를 구성하여 얼굴영상처리에 대한 대규모 프로젝트가 진행중이며 그간 개발되어 기술이 일부 실용화에 근접한 단계에 있다. 이런 상황에서 본 논문에서는 푸리에 변환과 다른 시간적, 주파수 해석이 가능한 웨이블릿 변화를 통해 얼굴 검출과 인식하는 알고리즘을 제시한다.

웨이블릿은 1983년 Moret에 의해 소개된 이후 푸리에 변환과 달리 시간 영역과 주파수 영역을 동시에 해석할수 있다는 장점을 가지고 신호를 분석하고 해석하는데 효과적인 수학적 도구로 알려져, 순수 수학기분야로부터 지표면 분석 영상처리 및 음성처리 같은 신호처리등 폭넓게 연구되고 있다. Wavelet 변화는 푸리에 변환에 기반을 둔 기존의 신호처리 알고리즘에 비해 속도가 빠르고 시간과 주파수 영역에서 신호의 국소화를 효율적으로 구현하기 때문에,

최근 신호 및 영상처리 분야영상개선 및 에지검출 기법, 영상재생, 영상압축등에서 많이 응용되고 있고 얼굴 검출 및 인식분야에서도 많은 연구가 이루어지고 있다.

본 논문에서는 얼굴 검출 알고리즘과 얼굴 인식 알고리즘으로 크게 두가지로 나누어서 제시한다. 입력 영상은 그레이 영상으로 웨이블릿 변환된 영상을 가지고 형판을 생성한후 얼굴 영역을 설정하고 그 후 수평, 수직 방향으로 투영시켜 눈의 위치 정보를 얻게 된다. 이 정보를 바탕으로 타원 정보를 이용하여 최종 얼굴영역을 검출한다.

그리고 얼굴 인식 알고리즘은 웨이블릿 패킷 2단계를 통해 생성된 16개의 부영상을 가지고 얼굴의 정보가 비교적 많은 저주파 영상에서는 수평, 수직 방향으로 투영시킨후 눈, 코 그리고 입의 기준선을 가지고 얼굴을 세가지로 분리하여 각 평균값과 분산값을 가진 특징 벡터를 추출하고 나머지 방향 성분을 가진 영상에서는 전체 영상에서 평균값과 분산값을 가진 총 18개의 특징 벡터를 추출한다. 그리고 유클리디언 거리를 이용해 특징벡터를 분류하도록 한다.

실험은 Pentium III 112MB 컴퓨터를 사용하여, 얼굴 인식 알고리즘만을 실험하였다. 배경이 없는 데이터를 이용하였는데, 200명의 각 개인당 3가지씩 총 600개의 영상 MIT FACES 데이터와 155명의 각 개인당 2가지씩 총 310장의 영상 FERET 데이터를 사용하여 P.J. Phillips가 FERET 실험과정으로 제안한 방법으로 수행 하였다. 성능은 얼굴 인식 방법에 기본적으로 사용하고 있는 PCA방법으로 수행한 결과와 비교하였으며, 제안한 알고리즘은 얼굴의 크기가 정규화되어 있지 않는 과정에서도 비교적 인식율이 높았으며, 고유 얼굴의 계산과 저장과정이 없다는 장점을 가지고 있었다.

---

**핵심되는말:** 얼굴검출, 얼굴인식, 웨이블릿 변환, 타원형 정보, 생체인식

# 제 1 장 서 론

---

과거에서 현재까지 과학과 공학적인 발전은 인간의 편의성 위주로 발전되었다. 발전 방향은 간편하고 편리하게 운용 할 수 있는 여건을 조성해 주면서 사용자의 요구에 수용되어진다. 또한 이런 요구들의 수용과 더불어 전반적인 기술 발전의 급변에 따라 편의성과 함께 중요하게 대두되어진 문제가 보안성이다. 기존 열쇠나 자물쇠, 신분증, 도장, 아이디 카드, 사인 등 물리적인 매개체로서 의존 되어왔던 보안 방법은 공간적인 확대에 따라 분실, 도난, 위조등의 문제도 커지게 되어 물리적 매체가 아닌 다른 방법을 통한 보안 문제의 해결이 필요하게 되었다. 부가적으로 물질적인 매체의 보관 필요성을 요구하지 않으면서, 개인을 식별하는 방법으로 생체 인식을 들수 있다.

생체 인식에는 얼굴 인식 이외에도 지문인식, 사인의 글자체 인식, 홍채인식, 자세나 걸음걸이 인식, 음성인식, 혈관인식 등이 있으며, 각각의 인식 및 검출의 특이성으로 인해서 약간의 제약을 가지고 있다. 그러나 다년간 연구 성과로 인해서 한가지 정보만이 아닌 다중 정보의 분석을 통한 인식이 가능해져 상용화의 안정성을 높이하고자 현재 연구가 진행되고 있다. 이미 생체 인식시스템 중 지문인식 시스템은 생산, 유지비용의 하락 및 현실적인 필요에 따라 일부 아파트 건설과 현금 자동인출기 등에서 실용화된 상태이다. 이밖에 건물의 보안을 위한 홍채 인식 시스템 또한 실용화된 상태이다.

한편 범죄자들의 걸음걸이등 습관적인 정보를 가지고 식별하는 방식이 공항의 입출국에서 시범적으로 사용되고 있다. 이중에서 사람의 얼굴이 개인마다 다르고, 원거리에서 식별이 가능하다는 장점을 가지고 있는 얼굴 검출의 방법이 다양한 방법을 통해 검출 및 인식상에 오류율이 적으면서 정확하도록 빠른 속도로 발전하고 있다.

정지 영상이나 동영상으로부터 다수의 얼굴을 인식하거나 한 개인의 얼굴을 식별하는 기술은 영상 처리나 패턴 인식, 컴퓨터 비전과 신경망 등과 같은 다양한 분야에서 활발히 연구되어지고 있으며, 상업적 많이 응용 되고 있다. 이러한 얼굴 인식 기술(Face Recognition Technology, FRT)은 여권, 신용 카드, 주민 등록증, 운전 면허증, 그리고 범인의 얼굴 사진 등과 같은 제한된 형태의 정지 영상을 사용하는 방법에서부터 동영상 감시 시스템 등과 같은 실시간 인식에 이르기까지 그 종류가 다양하다.

일반적으로 주어진 배경의 정지 영상이나 동영상에 대하여 입력 영상 내의 한 명 또는 그 이상의 사람이, 보관되어진 데이터베이스 내에 존재하는 사람인지 확인하는 것으로 정의될 수 있으며, 빠르고 정확한 탐색을 위해서 인종, 나이, 성별 등과 같은 부수적인 정보들이 사용될 수도 있다. 일반적으로 얼굴 인식의 문제는 대부분 영상으로부터의 얼굴 영역의 분리, 얼굴 특징점 추출, 그리고 분류 과정의 세 부분으로 나누어 처리되어진다[1].

인간과 기계의 얼굴 인식에 관한 수많은 연구들은 과거 20년이 넘는 기간 동안 심리학, 신경 과학, 그리고 공학 분야에서 폭넓게 진행되어 왔다. 얼굴 인식 기술의 초기 단계인 1970년 중반 동안에는 입력 얼굴로부터 특징점을 추출하여 저장되어 있는 특징점과 비교하는 전형적인 패턴 분류 방법이 주류를 이루었다. 하지만 많은 제약점과 낮은 성능으로 인해 1980년 대에서는 거의 연구가 되어지지 않았다. 그러나, 휴먼 인터페이스 중심의 연구가 활발해졌던 1990년 대에 이르러서는 이 얼굴 인식 기술이 매우 중요한 이슈로 성장하였다.

특히 최근 몇 년 동안 정면 얼굴 영상을 사용하는 인식 방법에 있어서 얼굴 영역의 분리와 눈, 코, 입 등과 같은 특징을 추출하는 연구에 많은 진전이 있었고, 영상의 통계적 처리 방법이나 퍼지 또는 신경망 등과 같은 여러 지능 알고리즘에 의한 접근 방법 등이 개발되었다[2][3].

정지 영상을 사용한 얼굴 인식 방법은 몇 가지 장단점을 가지고 있다. 예를 들어 여러 범죄자의 사진 중에서 특정인의 얼굴을 찾아내는 문제의 경우 사진 촬영 시 가해지는 여러 가지 제약 조건들에 의해서 얼굴의 추출이 다소 쉬워질 수 있으나, 공공 장소나 사무실과 같은 복잡한 배경에 인물이 들어 있는 사진이나 정지 영상에서는 얼굴의 추출이 어렵다[1]. 한편 감시 카메라로부터 얻어진 동영상의 분석에 있어서는 사람의 움직임을 고려하여 보다 쉽게 얼굴을 분리해 낸 연구 결과로 발표된 바 있다[2].

얼굴 검출 및 인식 분야는 영상신호처리, 영상 분할, 신경회로망 또는 통계적 패턴인식 기술등을 종합하는 기술로서 선진 외국의 경우, 기업은 물론 정부 심지어 국제 협력기구의 주도하에 1970년대부터 학제적 공동연구가 본격 전개되기 시작하였으며 현대 유럽의 경우 여러 국가에서 공동으로 연구 단체를 구성하여 얼굴영상처리에 대한 대규모 프로젝트가 진행중이며 그간 개발된 기술이 일부 실용화에 근접한 단계에 있다.

1990년대부터 시작된 연구의 방향은 검출이 비교적 용이하고 시간적으로 효율성이 높은 특징점을 기반으로 한 연구와 다중 얼굴 검출 및 기타 요건에 의해 검출이 곤란한 경우에 적용되어지는 영상을 기반으로 한 연구가 진행되고 있는데, 두 가지 방법을 모두 이용하여 간단하고 검출의 효율 및 정확도를 높이기 위한 연구 또한 진행되고 있다.

본 논문에서는 얼굴의 특징인 눈의 위치를 웨이블릿 변환된 영역을 적분투영시키는 방법을 통해 파악하고 단순한 타원형 식을 가지고 얼굴 영역을 검출하는 방법을 제시한다.

그리고 검출과 별도로 MIT 데이터베이스와 FERET 데이터베이스에서 웨이블릿 패킷 변환하여 얻어진 영상들의 평균과 분산값을 가지고 벡터를 추출한 후 유클리디언 거리 측정법을 사용하여 인식한다.

본 논문의 구성은 다음과 같다. 2장에서는 현재까지 연구되어진 특징점과 영상을 기반으로한 얼굴 검출과 얼굴 인식 방법들에 대해 설명하고, 3장에서는 얼굴 검출 및 인식 기법인 타원형 정보를 통한 방법과 웨이블릿 변환 그리고 비교 설명되어질 주성분 분석(PCA)를 통한 고유 얼굴(Eigenface) 방법에 대해 설명한다. 그리고 4장에서는 얼굴 검출 및 인식 알고리즘을 제시하고 마지막으로 5장에서는 기존 PCA방법과의 비교와 제안한 방법에 실험 결과 그리고 마지막으로 6장에서 결론 및 추후 과제를 제시한다.



## 제 2 장 기존의 얼굴 검출 및 인식 방법

---

오늘날 컴퓨터 기술의 발전은 인간생활이 인공지능에 의해서 향상되는 진보된 기계적 세계로 발전하는 기틀이 되었다. 이러한 과정은 기계에 지능을 접목하고자 하는 시도로 발전이 촉진 되었다고 할수있다. 예를 들어서 컴퓨터 비전은 인간의 시각을 복제하려는 목적을 가지고 연구 되었다. 전통적으로 컴퓨터 비전 시스템은 시각을 이용한 조립 라인 검사의 지루하고 반복적인 작업들을 수행하는 것과 같이 특정한 작업들에서 사용되어왔다. 이 분야에서의 현재 발전은 얼굴인식과 비디오 코딩 기법들과 같이 더 일반화된 비전 응용에 대한 연구가 활발히 진행되고 있다.

### 2.1. 얼굴 검출 및 인식 방법들

현재 많은 얼굴인식 기법들은 비슷한 크기들의 전면 얼굴들(frontal faces)의 이용가능성을 가정한다. 그러나 실제로 이러한 가정은 얼굴외양과 환경조건들의 변화되는 성질 때문에 많은 오류를 가지고 있다. 그림 2-1에 나타난 영상들은 얼굴 분류(face classification) 연구에서 사용되는 일반적인 테스트 이미지들이다. 그림 2-2에 있는 영상과 같이 현실적으로 얼굴은 복잡한 배경에서 그리고 많은 다른 위치들에서 나타날 수 있다. 배경의 어떤 영역들을 얼굴로서 잘못 인식될 수 있다. 이러한 문제를 제거하기 위해서는 배경으로부터 얼굴영역을 국소화하여 추출해야 한다..



그림 2-1. 얼굴인식을 위해서 훈련에 사용되는 전형적인 이미지들

얼굴 검출 및 인식은 사람에게는 힘들이지 않고 수행할 수 있는 시각적 작업들 중 하나이지만 컴퓨터로서의 관점은 영상이 입력시 알려지지 않은 수의 얼굴들의 위치를 결정하고 조명, 방향 그리고 카메라 거리와 무관하게 배경으로부터의 얼굴들 및 얼굴 특징들의 분할, 추출하는 것은 쉽지 않다.



그림 2-2 실제로 적용할 얼굴 영상

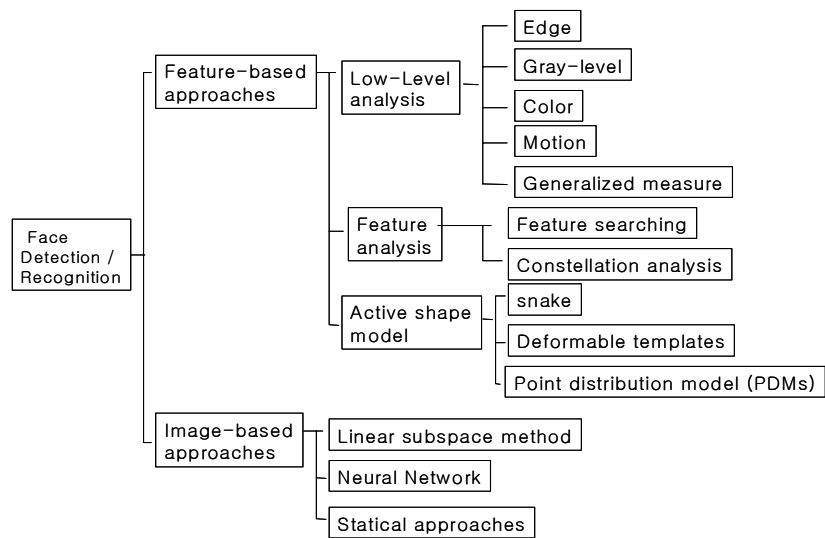


그림 2-3. 방법들에 따라서 나누어진 얼굴검출 및 인식방법

1970년대초 얼굴검출 및 인식은 비교적 간단한 기법들이 사용되었다. 이들 기법들은 단순한 배경, 정면 얼굴과 같은 상당한 제약을 필요로 하여 현실적인 응용이 어려웠다. 이들 문제점들을 극복하기 위해서 연구 발전은 실제적인 얼굴인식 및 비디오 코딩 시스템들이 현실이 되기 시작한 1990년대가 될 때 까지 정체되었다. 그러나 움직임, 컬러 그리고 일반화된 정보의 사용과 함께 통계학 및 신경망의 사용은 다른 위치에서 카메라에 찍힌 혼잡한 장면들로부터 얼굴들을 검출하는 것을 가능하게 하였다. 그리고 얼굴 특징들의 위치를 정확하게 결정하고 추적할 수 있는 변형 가능한 템플릿들 및 활성 윤곽선들과 같은 특징 추출기들의 설계에서 많은 진전이 이루어졌다. 얼굴검출 및 인식기법들은 얼굴 정보를 이용하는데 있어서 다음 두 가지 방법으로 나누어 볼수 있다.

첫 번째 방법은 얼굴의 특징들을 사용하는 고전적 검출방법론이다. 이 경우 피부색깔과 얼굴형태와 같은 얼굴의 외양 특성들이 이용되는데, 일반적으로 이들 기법들에서 얼굴검출 방법들은 정면으로부터 얻어진 시각적 특징들의 거리, 각도들 그리고 면적을 측정함으로써 수행된다. 특징들은 주요한 구성요소들이기 때문에 이들 기법들은 특징 기반의 방법이라 할수 있다. 이들 방법들은 1970년대 초에 시작한 얼굴검출 연구의 주된 방법이다.

그리고 두 번째 방법은 패턴인식 이론을 접목한 영상기반 방법이다. 현재 발전들의 장점을 취하는 기법들은 얼굴검출을 일반적인 인식문제로서 다루는데, 이 방법은 얼굴들의 영상 기반 표현들을 통해 특징추출 및 분석 없이 훈련 알고리즘들을 사용하여 얼굴을 몇가지 그룹으로 직접 분류한다. 첫 번째 분류인 특징기반의 방법과 다르게 상대적으로 새로운 기법로서 매핑과 훈련체제를 통해서 얼굴 정보가 시스템에 이용되는 방식이다.

## 2.1.1 특징 기반 연구

특징 기반 방법의 발전은 세 가지 분야들로 나누어 볼수 있는데, 혼잡한 배경에서 얼굴의 위치를 결정하는 전형적인 얼굴검출 문제가 주어지면 먼저 저차원 분석이 그레이 스케일 및 컬러와 같은 픽셀 특성들을 사용하여 시각적 특징들의 분할을 다루지만 저차원 본질적인 특성 때문에 이 분석으로부터 생성되는 특징들은 분명하지 않다.

저차원 특징중 모서리 특징은 안면 특징들의 위치를 결정할 목적으로 사진들로부터 얼굴들의 외곽선을 분석하는 것에 기초한 Sakai가 초기의 얼굴검출 연구에서 적용되었다[5]. 사람의 머리 윤곽을 추적하기 위해서 Sakai의 연구에 기초하여 계층적 프레임워크를 Craw[4] 제안하였으며, 안경을 검출하기 위해서도 적용되었다.[6.7]

얼굴 내부의 그레이 정보는 눈썹, 눈동자 그리고 입술들과 같은 얼굴 특징들은 일반적으로 피부와 같은 얼굴 영역들보다 더 어둡게 나타나므로 다양한 얼굴 부분들을 구별할수 있다. 먼저 콘트라스트 확장과 그레이 스케일 모폴로지 루틴들을 이용하고, 검출은 저수준 그레이 스케일 임계값을 이용한다. 컬러 해석으로부터 간접적으로 얻어지는 얼굴 후보들 내부에서 어두운 안면 영역들을 탐색하는 로봇을 구현 하기 위해 한 쌍의 눈의 위치들을 결정하기 위해서 가중된 사람 눈 템플릿을 사용한 Wong[8]. 여덟 개의 어두운 근방들에 의해서 둘러싸이는 밝은 픽셀에 의해서 정의되는 국소 최대값이 코 끝과 같이 밝은 안면 점들을 가리키기 위해서 사용한 Hoogenboom과 Lew[9]등이 있다.

컬러는 객체 외양을 구분에 사용되는데, 적색, 녹색 및 청색 의 주 색상 성분들의 조합들에 의해 정의되는 RGB 표현이다. 피부외양에서의 주요한 변화는 대개 휘도 변화에 기인하기 때문에[10] 일반적으로 정규화된 RGB 컬러들이 적용된다.

비디오 시퀀스에 적용되는 움직임 정보는 움직이는 객체들의 위치를 결정하는데 편리한 수단으로 얼굴영역 이외에도 안면 특징들의 위치를 결정하기 위해서 프레임 차를 채택하는 기법들을 Luthon & Lievin[11]과 Low[12]등이 제안하였다. 그리고 Reisfeld 와 Yeshurun은 모서리 픽셀 연산들에 기초하는 일반화된 대칭 연산자를 이용하였다[13].

특징 분석에서 시각적인 특징들은 얼굴의 형태적인 정보로서 얼굴 및 안면 특징들을 이용한다. 특징분석을 통해 특징들의 모호함들이 감소되고 얼굴 및 안면 특징들의 위치가 결정되는데, 머리와 어깨 영역에 있는 작은 영역을 얼굴이라 가정하고 얼굴 영역에서 한 쌍의 어두운 영역을 눈으로 표현하여 얼굴을 검출한다. 한 쌍의 눈은 나란한 특징을 가지고 있기 때문에 많이 이용되는 방법으로 De Silva et al.에 의한 안면 특징 추출 알고리즘은 특징탐색의 좋은 예이다[14].

고정된 조건들 하에서 모델링된 다양한 얼굴 이미지들로부터 얻어진 경험적 정보에 의존성이 높는데, 경직된 본성 때문에 검출이 곤란하다. 통계학적 해석과 같은 더 강인한 모델링 방법들을 사용하여 얼굴 같은 배열들은 안면 특징들을 그룹화함으로써 문제점을 다루고 있다. 다중 스케일 가우시안 미분필터로부터 검출된 특징들 위에서 통계학적 형상이론(statistical shape theory)을 사용한 Burl et al [15]이 있다.

마지막인 활성 형상 모델들(Active Shape Models)은 1980년대 말에 제안된 스네이크 알고리즘에서 부터 최근의 점 분산된 모델들(Point Distributed Models)에 이르기까지 모델들은 눈동자 및 입술 추적과 같은 복잡하고 비 경직된 특징 추출을 목적으로 개발되었다.

활성 형상 모델은 모서리들, 밝기와 같은 국소 이미지 특징들과 상호작용하고 점차적으로 특징의 형상을 얻기 위해서 변형되는데, Kass에 의해서 1987년에 처음 소개된 스네이크라 불리는 일반적인 능동 윤곽을 사용한다[16].

스네이크는 또한 Saito에 의해서 안경검출을 위하여 사용되었다[19]. 그 뒤에 얼굴 특징들의 우선성을 고려하고 스네이크의 성능을 높이기 위해서 변형 가능한 11개의 파라미터를 이용한 눈 템플릿을 Yuille에 의해서 소개되었다[17]. 그 후에 Cootes는 사람 얼굴의 효율적인 해석을 제공하기 위해서 통계학에 기초한 형상의 간결한 매개변수화된 기술인 PDM(Point Distributed Models)인 일반적이고 유연한 모델의 사용을 제안하였다[18]. Cootes의 모델은 훈련절차에 따라서 형상들로만 변하는 것이 허용되는 라벨이 붙은 점들의 집합으로 표현한다.

### 2.1.2. 영상 기반 방법

특징기반 방법은 얼굴외양과 환경조건들의 변화시에 얼굴검출이 곤란하다. 다중 얼굴들을 검출하는 것과 같은 더 어려운 과정에서 수행할 수 있는 필요성으로 인해 얼굴검출을 패턴인식 문제로서 다루었는데 훈련과정으로 얼굴 패턴을 인식하기 위해 학습화를 수행함으로써 불완전하거나 부정확한 얼굴영역 검출 에러를 제거한다. 얼굴패턴들을 인식하는 기본적인 방법은 얼굴과 얼굴이 아닌 클래스들로 분류하는 훈련 과정이다. 영상 기반 방법은 2차원적으로 표현한다.

영상 기반 방법들 대부분은 얼굴들을 검출하기 위하여 윈도우 스캐닝 기법을 적용한다. 윈도우 스캐닝 알고리즘은 본질적으로 모든 스케일들에서 가능한 얼굴 위치들을 위한 입력 이미지의 자세한 탐색이지만 거의 모든 이미지 기반 시스템들을 위해 알고리즘의 구현에서의 변형된다. 전형적으로 스캐닝 윈도우의 크기, 하위샘플링 비율, 스텝 크기, 그리고 반복횟수들은 제안되는 방법과 계산적으로 효율적인 시스템을 위한 필요성에 의존하여 변한다.

영상 기반 방법들을 선형 부분공간 방법들(Linear Subspace Methods), 신경망(Neural Networks), 그리고 통계학적 접근법등이 있다.

선형 부분공간 방법은 1980년대 말에 Sirovich와 Kirby[20]은 사람 얼굴들을 효과적으로 표현하기 위해서 주성분 해석(Principal Component Analysis)을 제안하였다. PCA는 공분산 행렬의 고유벡터들에 의하여 표현한 얼굴들의 주성분들을 찾는 방법이며, 얼굴집합에 있는 각 개별 얼굴이 적절한 가중들을 사용하여 더 일반적으로는 고유 얼굴들(eigenfaces)로 나타내어 선형조합으로 나타내는 방법이다. Turk와 Pentland는 얼굴인식을 위하여 이 기법을 발전시켰다[21].

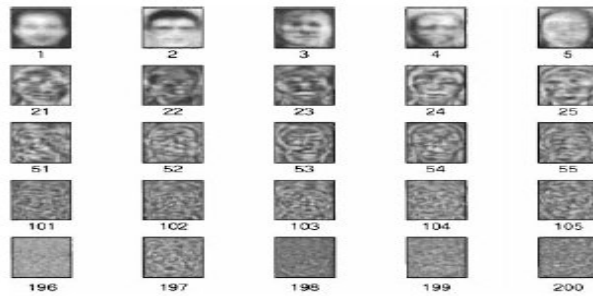


그림 2-4 ORL 데이터 베이스의 고유얼굴 예  
( 아래 숫자는 고유값에 따른 주성분 수 )

얼굴 및 안면 특징 검출을 위하여 Fisher의 선형판별식[22]과 조합되어 사용하여 훈련 집합에 있는 다양한 안면 특징 템플릿들로부터 얻어진 고유특징들(고유 눈들, 고유 코, 고유 입)로부터 안면특징 검출기를 제안한 Pentland et al.[23], 확률론 틀 내에서 이 기법을 더 발전시킨 Moghaddam과 Pentland[24], 컬러, 3D 및 움직임 정보에 기초한 Jebara 및 Pentland[25], 그리고 얼굴들 및 배경 클러스터 모두를 모델링하기 위해서 사용한 Meng 및 Nguyen[26]등이 있다.

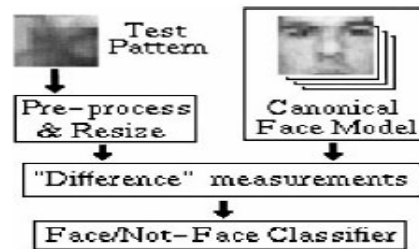


그림 2-5. Sung 및 Poggio의 시스템에서의 분류

PCA는 많은 경우들에 객체 클래스를 표현하기 위하여 부분공간을 생성하는 직관적이고 적절한 방법이다. 그러나 얼굴 이미지들의 다양체를 모델링하기 위하여 PCA는 반드시 최적인 것은 아니다. 얼굴공간은 그것을 부분클래스들로 분할함으로써 더 잘 표현될 수 있으며 그 결과 몇 가지 방법들이 이것을 위하여 제안한 바 있으며 그 중 대부분은 몇몇 다차원 가우시안들의 혼합에 기초한다. 이 기법은 최초로 Sung과 Poggio에 의해서 얼굴검출을 위하여 적용되었다.

신경망을 이용한 방법은 얼굴검출을 포함하여 패턴인식 문제들을 위해서 널리 사용되는 기법이다. 얼굴인식을 위한 다소 기초적인 신경망 방법들에 대한 소개는 Viennet 및 Fougelman Soulie에서 찾아볼 수 있다[28].

Rowley et al[29]의 고급 뉴럴 기법은 그림 2-6에서 보이는 눈의 망막에 연결된 신경망에 얼굴지식을 결합하여 수행된 연구였다. 이 방법의 문제점은 윈도우 스캐닝 기법들을 사용할 때 중복되어 검출되는 것이다.



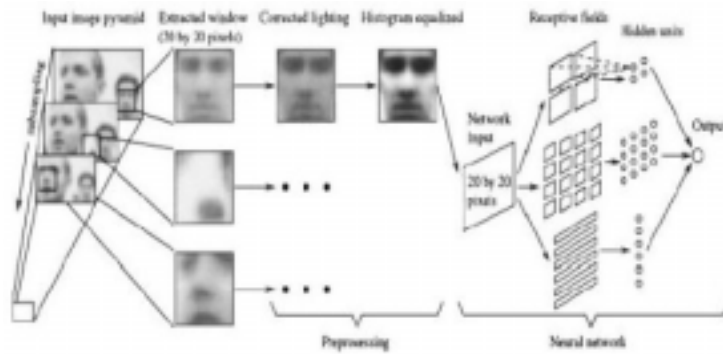


그림 2-6 Rowley et al.의 시스템

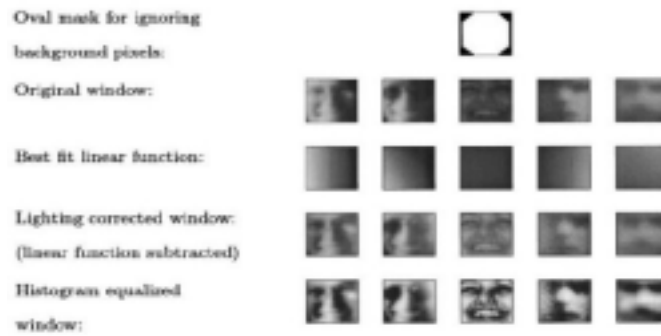


그림 2-7 Rowley et al.과 Sung 및 Poggio에 의해 적용된 전처리 방법

Feraud[30]는 제약된 생성 모델(Constrained Generative Model)에 기초하여 세 개층으로 표현한 뉴럴 기법이 제안하였다. Lin은 완전히 자동적인 얼굴인식 시스템이 확률론적 결정 기반의 신경망(Probabilistic Decision Based Neural Networks)에 기초한 방법을 제안하였다[31]. Roth et al.[32]은 얼굴들과 비얼굴들의 클래스들을 표현하는 두 개의 선형 임계 단위 (Linear Threshold Units)로 구성되는 신경망인 SNoW(sparse network of windows)라 불리는 새로운 학습 아키텍처가 얼굴검출을 위하여 적용된다.

선형 부분공간법들 및 신경망들과는 별도로 얼굴검출을 위한 몇 가지 다른 통계적 방법들이 존재하는데, 정보이론, Support Vector Machine 그리고 Bayes의 결정규칙등이 있다.

Osuna et al.에서 Support Vector Machine(SVM)[34]가 얼굴검출을 위하여 적용되었다[33]. Kumar 및 Poggio은 최근에 얼굴들의 실시간 추적과 분석을 위하여 Osuna et al.의 SVM 알고리즘을 사용하였다.

## 제 3 장 얼굴 검출 및 인식 기법

---

본장에서는 제안하는 각기 다른 형태의 얼굴 검출 및 인식기법의 대해 살펴 보도록 한다.

### 3.1 타원정보로 얼굴영역 검출기법.

얼굴의 일반적인 형태가 타원임을 알고 있으므로 그 정보를 이용하여 얼굴 검출을 수행한다. Sirohey[39]의 얼굴 검출기법은 인간의 머리 부분의 타원형에 가장 적합한 에지들을 찾아내어 배경으로부터 얼굴을 검출하는 것이다. 타원형의 물체 검출에 가장 많이 사용되었던 것은 Hough변환을 이용한다[40,41]. 하지만 이 방법은 배경이 깨끗한 경우에도 엄청난 계산량을 요구하는 단점이 있으므로, 낮은 해상도에서 높은 해상도의 영상까지에서 적응적으로 타원 파라미터를 찾아내는 적응적 Hough변환이 이용된다[42,43]. 하지만 이 방법 또한 계산량이 많아서 Sirohey는 우선 배경의 에지와 얼굴의 에지를 분리해 내고 선형화된 타원 방정식을 이용하여 얼굴 타원을 검출하는 방법을 이용하였다.

#### 3.1.1 에지 검출과 에지 연결과정.

타원형 에지를 찾기 위한 첫 단계로서 데이터 베이스내의 얼굴 영상에서 Canny 에지 검출기를 이용하여 에지를 찾아 낸다. 이렇게 구한 에지 영상에서는 수많은 에지간의 교차점들이 존재한다. 얼굴 부분과 배경부분의 에지를 구분하기 위해서, 두 개의 교차하는 에지가 서로 다른 물체에서 나온 것이라면 이러한 교차점  $\epsilon(x,y)$ 이라 정의하고 그 주변  $3 \times 3$  픽셀을 없애게 된다. 이런 과정을 거치면 한물체에 해당하는 에지 부분이 떨어져 있던 에지가 여러개로 분리되는 문제가 발생한다. 그러므로 이런 부분을 다시 연결해 주는 과정이 필요하게 된다.

에지 연결과정은 우선 조각난 에지 조각들에 번호를 매기는 것을 시작으로 한다. 다음에는 앞서 찾아낸 모든 교차점  $\epsilon(x, y)$ 들의 주변에서 어떤 에지 부분이 같은 물체에 속하는지는 판단한다. 교차점  $\epsilon(x, y)$ 주변에 세 개의 에지가 존재한다고 가정한다. 그리고,  $\epsilon(x, y)$ 로부터 n픽셀만큼 떨어진 i 번째 에지 조각위의 에지 픽셀을  $\epsilon_i(x, y)$ ,  $i \in \{1, 2, 3\}$ 이라 하고 백터를 정의하는 식(3.1)이다..

$$v_i(\epsilon(x, y) - \epsilon_i(x', y')) \quad \forall i \in \{1, 2, 3\} \quad (3.1)$$

여기서  $(x', y')$ 는  $\epsilon(x, y)$ 로부터 n 픽셀 떨어진 에지 픽셀의 위치를 나타낸다. 각각에 대해 단위백터  $e_i, \forall i \in \{1, 2, 3\}$ 을 구하면 식 (3.2)가 되고 서로의 내적을 구하면 식(3.3)과 같다.

$$e_i = \frac{v_i}{\|v_i\|}, \quad \forall i \in \{1, 2, 3\} \quad (3.2)$$

$$e_i \cdot e_j = \cos \theta_{ij}, \quad \forall i, j \in i \neq j \quad i, j \in \{1, 2, 3\} \quad (3.3)$$

3개의 내적값중에서 최소의 값을 갖는 두개의 에지들이 동일한 물체의 경계선 이라고 볼수 있다. 그러므로 이러한 두 개의 에지들에 다음과 같이 동일한 번호를 매길수 있다.

$$L_i = L_j$$

$$\text{if } \cos \theta_{ij} = \min(e_k \cdot e_l), \quad k, l \in \{1, 2, 3\} \quad (3.4)$$

이렇게 각 격자점마다 다시 연결된 에지 영상을 바탕으로 얼굴을 나타내는 타원형을 찾게 된다.

### 3.1.2 타원형 찾기

앞에서 얻어진 에지 영상으로부터 타원형에 맞는 에지들을 찾아내면 그것으로 얼굴 검출이 이루어지게 된다. 우선 모든 에지들에 번호가 주어졌으므로 한쌍의 에지에 대해서 그것이 타원형에 적합한지 검사한다. 이를 위해 우선 타원의 식을 선형화한다. 일반적인 타원의 식 (3.5)에  $a^2$ 을 곱해서 선형화한 식을 식(3.6)에 나타내었다.

$$\frac{(x_i - x_0)^2}{a^2} + \frac{(y_i - y_0)^2}{b^2} = 1, \quad i \in \{1, 2, \dots, N\} \quad (3.5)$$

$$2x_i a_0 - y_i^2 a_1 + 2y_i a_2 - a_3 = x_i^2, \quad i \in \{1, 2, \dots, N\} \quad (3.6)$$

여기서 N은 특정번호를 가진 에지 픽셀의 수를 나타낸다.

각 파라미터들은 다음식 (3.7)과 같다.

$$a_0 = x_0, \quad a_1 = \frac{a^2}{b^2}, \quad a_2 = \frac{a^2}{b^2} y_0, \quad a_3 = x_0^2 + \frac{a^2}{b^2} y_0^2 - a^2 \quad (3.7)$$

식(3.6)은 각 파라미터들에 대해 선형적인 특성을 갖는 선형 방정식이다. 이제  $a_0$ 에서  $a_3$ 까지 구하면 타원의 중심과 장축, 단축을 구할수 있다. 각 번호의 에지에 대해서 이 파라미터들은 다음식(3.8)과 같이 구할수 있다.

$$\begin{bmatrix} 2x_1 & -y_1^2 & 2y_1 & -1 \\ 2x_2 & -y_2^2 & 2y_2 & -1 \\ \vdots & \vdots & \vdots & \vdots \\ 2x_N & -y_N^2 & 2y_N & -1 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} x_1^2 \\ x_2^2 \\ \vdots \\ x_N^2 \end{bmatrix} \quad (3.8)$$

식(3.8)은  $AX=C$  형태로서  $A$ 는  $N \times 4$ ,  $X$ 는  $4 \times 1$ ,  $C$ 는  $N \times 1$ 이다.  $X$ 를 구하기 위해서 의사 역행렬을 식 (3.9)에서와 같이 이용할수 있으며 구한식은 (3.10), (3.11)과 같다.

$$X=(A^tA)^{-1}A^tC \quad (3.9)$$

$$A^tA = \begin{bmatrix} 4 \sum_{i=1}^N x_i^2 & -2 \sum_{i=1}^N x_i y_i^2 & 4 \sum_{i=1}^N x_i y_i & -2 \sum_{i=1}^N x_i \\ -2 \sum_{i=1}^N x_i y_i^2 & \sum_{i=1}^N y_i^4 & -2 \sum_{i=1}^N y_i^3 & \sum_{i=1}^N y_i^2 \\ 4 \sum_{i=1}^N x_i y_i & -2 \sum_{i=1}^N x_i y_i^3 & 4 \sum_{i=1}^N y_i^2 & -2 \sum_{i=1}^N y_i \\ -2 \sum_{i=1}^N x_i & \sum_{i=1}^N y_i^2 & -2 \sum_{i=1}^N y_i & N \end{bmatrix} \quad (3.10)$$

$$A^tC = \begin{bmatrix} 2 \sum_{i=1}^N x_i^3 \\ - \sum_{i=1}^N x_i^2 y_i^2 \\ 2 \sum_{i=1}^N x_i^2 y_i \\ - \sum_{i=1}^N x_i^2 \end{bmatrix} \quad (3.11)$$

이러한 과정은  $L_i$  와  $L_j$ 의 번호가 붙은 한쌍의 에지들에 대해 행해진다. 일단  $X$ 를 찾고 나면, 식(3.7)에 의해 그타원의 중심과 장축 단축을 찾아낼수 있다. 이렇게 찾아낸 일련의 파라미터들 중에 얼굴의 타원형에 해당하는 것들만 가려내야 한다. 우선 장축과 단축의 비율이 일정한 범위를 만족하는 파라미터들만 골라낸다. 이렇게 선택된 파라미터  $P_i=\{x_{0i}, y_{0i}, a_i, b_i\}$ 를 갖는 타원후보를  $C_\epsilon$ 이라 하면 식(3.12)와 같이  $C_\epsilon$ 에 대해 모든 에지 조각들의 오차를 구한다.

$$e = \frac{(x_i - x_0)^2}{a^2} + \frac{(y_i - y_0)^2}{b^2} - 1, \quad \forall i \in \{1, 2, \dots, M\}$$

$$\forall V_j, \quad j \in \{1, 2, \dots, M\} \quad (3.12)$$

즉  $j$ 의 번호가 붙은 모든 에지 픽셀  $(x_i, y_j)$ 에 대해 그 타원에 대한 오차를 구하여 만약  $e < \mu$  ( $\mu$ 는 문턱값)이면  $(x_i, y_j)$ 는 파라미터를 갖는 타원에 적합하다고 할수 있다. 각각의 번호가 붙은 에지에 대해서 이런 픽셀이 전체 픽셀수의 반 이상이면 이 번호의 에지는 그 타원에 속한다고 판단한다. 이렇게 선택된 타원 후보들 중에 그것에 속하는 에지의 픽셀수가 가장 많은 타원을 얼굴의 타원으로 결정한다.

이러한 방법은 영상의 대부분이 얼굴로 채워져 있지 않은 경우에서는 얼굴을 제외한 다른 에지들이 너무 많기 때문에 시간이 많이 걸리고 얼굴 검출이 어려우며 여러명의 얼굴이 있는 응용분야에서는 적용이 불가능하다.

### 3.2 PCA를 이용한 얼굴 인식 방법.

본 논문에서 제안하는 방법과 비교하기 위한 PCA( Principal Component Analysis ) 방법은 Matthew A. Turk 와 Alex P. Pentland가 1991년에 발표한 논문에서 소개된 것으로 이 방법은 인식 하고자 하는 얼굴 샘플들의 공분산 행렬의 고유 벡터를 계산하고, 새롭게 입력된 얼굴을 고유벡터 공간으로 투사시켜 그 성분들을 비교함으로써 인식을 수행하는 방법으로 그 과정은 다음과 같다.

①. 얼굴 영상의 크기가  $n \times n$ 이고, 훈련 집합에 있는 얼굴 영상의 개수가  $m$  개 일 때, 각각의 영상을  $n^2$ 개의 성분을 가지는 1차원 칼럼벡터들로 표현하면  $n^2 \times m$  행렬로 표현되고, 훈련집합은  $\Gamma_n, n=1, \dots, m$  이다.

여기서  $\Gamma_n$ 은 훈련 집합안에 있는 그림 3-1은 각각의 얼굴을 나타낸다.



그림 3-1 주성분 분석을 위해 훈련 집합안에 영상들

②. 그 후 훈련 집합 (Training set)의 평균 얼굴( $\phi$ )을 구하는데, 식 (3.13)은 평균 얼굴을 구할수 있는식이다..

$$\phi = \frac{1}{m} \sum_{k=1}^m \Gamma_k \quad (3.13)$$

평균얼굴은 다음 그림 3-2 이다.



그림 3-2 평균 얼굴



③. 원 영상에서 평균얼굴을 뺀  $\Phi_i = I_i - \psi$  벡터를 구한다. 원 영상에서 평균얼굴을 뺀 영상의 예를 아래 그림 3-3과 같다.



그림 3-3 원영상에서 평균을 뺀 영상

이렇게 얻어진  $\Phi_i$  라는 벡터를 이용해서 공분산 행렬(covariance)을 만드는데 식 (3.14)와 같다.

$$C = \frac{1}{m} \sum_{n=1}^m \Phi_n \Phi_n^T \quad (3.14)$$

이 공분산 행렬( $N^2 \times N^2$ )에서 데이터의 분포를 가장 잘 표현하는 M의 정규직교 벡터  $\mathbf{u}_n$ 와 이와 관련된 고유값  $\lambda_k$ 를 구한다. 여기서 벡터  $\mathbf{u}_k$ 와  $\lambda_k$ 를 각각 공분산 행렬의 고유벡터(eigenvectors)와 고유값(eigenvalues)이다.

고유값은 얼굴 영상의 특징을 나타내는데, 그 값의 유용성에 따라 크기 순으로 정렬된다. 이렇게 얻어진 고유벡터들을  $N \times N$ 으로 재배열하여 고유얼굴(eigenfaces)을 나타낸다.

이렇게 얻어진 고유얼굴들이 그림 3-4에 나타내었다.



그림 3-4 고유얼굴들

④. 고유얼굴을 이용한 인식과정은 다음과 같다. 일단 새로운 얼굴 영상이 입력되면 고유벡터 공간으로 투사시켜 그 성분 값(weights)을 구하여, 새로이 입력된 영상을 고유얼굴들로 표현한다. 고유얼굴(eigenfaces)을 기저함수로 사용하여 입력된 영상의 고유벡터 공간에서의 성분 값을 식 (3.15)를 통하여 얻는다.

$$\omega_k = u_k^T (Y - \Psi) \quad \text{for } k=1, \dots, M \quad (3.15)$$

이렇게 얻어진 성분 값(weights)은 입력 영상을 표현하는데 있어서 각각의 고유얼굴의 공헌도를 의미하는 벡터를 구한다. 벡터는 다음의 식(3.16)과 같다.

$$\Omega^T = [\omega_1 \omega_2 \dots \omega_M] \quad (3.16)$$

이 값이 구해지면 데이터 베이스에 있는 얼굴영상들의 고유벡터 공간에서의 성분벡터( $\Omega_k$ )와 비교해서 유클리디언 거리 ( Euclidian distance(  $\epsilon_k = \|\Omega - \Omega_k\|$  )가 최소가 되는 얼굴영상이 바로 입력영상과 가장 흡사한 얼굴로 인식이 된다. 입력영상에 대한 인식결과는 다음 그림 3-5와 3-6과 같다.



그림 3-5 입력영상



그림 3-6 인식된 얼굴

이 PCA방법은 다른 얼굴 인식 방법들에 비해 속도와 단순함 그리고 학습능력에 있어서 장점을 가지고 얼굴 영상에서의 작고 점진적인 변화에 상대적으로 덜 민감하게 사용되고 있다. 그러나 몇가지 제약점을 가지고 있는데, 얼굴 이미지의 다양체를 모델링하기 위해서는 얼굴 공간을 부분 클래스로서 분할함으로써 더 잘 표현되어질수 있기 때문에 최적의 방법은 아니다.

### 3.3 Wavelet 이론

웨이블렛 변환은 1983년 Moret에 의해 소개된 이후 푸리에 변환과 달리 시간 영역과 주파수 영역을 동시에 해석할 수 있다는 장점을 가지고 있으며, 신호를 분석하고 해석하는데 효과적인 수학적 도구로 알려져, 순수 수학분야로부터 지표면 분석 영상처리 및 음성처리 같은 신호처리등에 폭넓게 사용되고 있다. Wavelet 변환은 푸리에 변환에 기반을 둔 기존의 신호처리 알고리즘에 비해 속도가 빠르고 시간과 주파수 영역에서 신호의 국소화를 효율적으로 구현하기 때문에 최근 신호 및 영상처리 분야영상개선 및 에지검출 기법, 영상재생, 영상압축등에서 많이 응용되고 있고 얼굴 검출 및 인식분야에서도 많은 연구가 이루어지고 있다[45~47].

과거 10여년동안 연구를 통해 웨이블렛은 매우 보편화되고 있고, 새로운 관심으로 부상하고 있으며, 최근에 웨이블렛 기저와 웨이블렛 변환을 위한 효과적인 알고리즘이 구성되었다. 그 중 웨이블렛의 유용한 특징중의 하나는 몇 개의 기저가 존재하더라도 주어진 응용에 따라 좀더 적당한 형태의 한가지 기저를 선택할 수 있다는 점이다. 더욱이 다른 변환식은 빠르게 구성하기 위해서 로그 함수 ( $N \times \log_2(N)$ )를 이용하는 반면에 웨이블렛은 계산된 수에 선형적으로 처리가 가능하기 때문에 빠르게 처리되어진다. 그리고 웨이블렛 변환은 하드웨어 구성을 위해 이산화 웨이블렛 변환이 이용되기도 한다. 만약 인식하는 일이 실시간적으로 필요하다면 하드웨어적인 처리과정은 매우 중요하다. 웨이블렛을 이용한 압축 알고리즘은 1997년 시드니에서 열린 12th WGI회의 이후에 JPEG2000 코딩에 소개된 후 적용되고 있다.

웨이블렛 변환은 다중 분해 해석 (Multiresolution Analysis ) 분야에서 제안되어, 영상과 신호처리에 적용 발전되었다[48]. 웨이블렛 변환은 주어진 함수  $f(x)$ 의 신호와 웨이블렛 집합의 내부 적(inner product)에 의해 웨이블렛 계수를 생성하며, 이 계수들과 관련된 웨이블렛 함수  $\psi(x)$ 는 여러 등급의 근사화로 표현된다.

웨이블렛 변환의 기본 개념은 임의의 함수  $f(x)$ 를 시간, 주파수 공간에서 동시에 지역성을 갖는 웨이블렛 기저함수(base function)의 선형결합(superposition)으로 표현하는 것이다. 원형 웨이블렛 함수 식 (3.16)  $\Psi(x)$ 를 이용하여

$$\Psi_{b,a}(x) = |a|^{-1/2} \Psi\left(\frac{x-b}{a}\right) \quad (3.16)$$

와 같은 웨이블렛 변환식을 얻을 수 있다. 여기서  $a$ 는 스케일 변수(scale variable)이고,  $b$ 는 이동 변수(translation variable)이다.

$\Psi_{a,b}(x)$ 은 웨이블렛 함수로서  $L^2(R)$ 공간에 속한 요소이며,  $L^2(R)$ 공간내의 함수  $f(x)$ 는 식 (3.17)의 조건을 만족해야 한다.

$$\int_{-\infty}^{+\infty} |f(x)|^2 dx, \int_{-\infty}^{+\infty} |\Psi_{a,b}(x)| dx = 0 \quad (3.17)$$

웨이블렛 함수는 푸리에 변환은 주어진 공간내에서 같은 시간과 주파수 대역에 위치하며[51], 빠른 감소 특성과 진동 정도를 나타내는 일반화된 형태로 표현이 가능하다.

$$(WT)_{(d,m)} = |a|^{-1/2} \int_{-\infty}^{+\infty} f(x) \Psi\left(\frac{x-b}{a}\right) dx \quad (3.18)$$

식 (3.18)에서  $a, b$ 는 실수이고, 함수  $f(x)$ 는  $L^2(R)$  공간내의 함수이며, 기저함수의 선형조합으로 구성된 식 (3.19)와 같이 나타내어질수 있다[48]

$$f(x) = f_a(x) + \sum_b \sum_{a=a_0}^{\infty} \langle \Psi_a(x), f(x) \rangle \Psi_{a,b}(x) = \sum_k c_a(k) \Phi_{a,b}(x) + \sum_b \sum_{a=a_0}^{\infty} d_a(k) \Psi_{a,b}(x) \quad (3.19)$$

$a, b \in \mathbb{Z}$   $\Phi_{a_0,b}(x)$ 는 스케일 함수이고  $\Psi_{a,b}(x)$ 는 웨이블렛 함수이다.

식(3.19)에서  $a_0$ 값이 0이거나 음의 무한대 값을 갖는다면, 함수  $f(x)$ 는 웨이블릿 함수로만 나타낼 수 있다. 첫 번째 합은 저해상도의 근사값이고, 두 번째 합은 고해상도의 근사값을 갖는다. 식 (3.19)가 스케일 함수의 저대역 통과 필터와 웨이블릿 함수의 고대역 통과 필터로 분리가 가능하다. 만약  $\Phi_{a_0, b}(x)$ 와  $\Psi_{a, b}(x)$ 가 어떤  $a$ 나  $b$ 에서 직교성을 갖는다면  $c_a(k)$ ,  $d_a(k)$ 는 다음 식(3.20)으로 나타내어진다..

$$c_a = \langle f(t), \Phi_{a, b}(t) \rangle = \int f(t) \Phi_{a, b}(t) dt$$

$$d_a = \langle f(t), \Psi_{a, b}(t) \rangle = \int f(t) \Psi_{a, b}(t) dt \quad (3.20)$$

식(3.20)에 의해 구해진 계수를 웨이블릿 계수이고, 이 식에서 다중 분해식은 시간-크기를 갖는 신호로 나타내어 질수 있으므로 식(3.16)의  $\Psi_{a, b}(x)$ 는 다음과 같이 일반화된다.

$$\Psi_{a, b}(x) = |a|^{-1/2} \Psi\left(\frac{x-b}{a}\right) \quad (3.21)$$

식(3.21)에 의해 식(3.19)는 다음과 같다

$$f(x) = \sum_k c_a(k) \Phi_{a, b}(x) + \sum_b \sum_{a=a_0}^{\infty} d_a(k) |a|^{-1/2} \Psi\left(\frac{x-b}{a}\right) \quad (3.22)$$

여기서,  $a$ 는 크기 변화 인자이고,  $b$ 는 위치 변화 인자 이다.

스케일 함수의 확장 인자와 변형인자는  $L^2(\mathbb{R})$  공간의 다중 분해 해석 방식에 의해 식 (3.23)과 같이 폐 부분공간으로 나타낼수 있다.

$$V_0 \subset V_1 \subset V_2 \subset \dots \subset L^2 \quad (3.23)$$

즉  $\bigcap_d V_d = V_d \setminus \{0\}$ ,  $close\left\{\bigcap_d V_d\right\} = L^2(\mathbb{R})$ 이다.

여기서,  $V_d$ 는 식(3.21)에 의해 확장된 부공간이다.

또한 식 (3.23)을 확장된 다른 부공간을  $W_1$ 이라 하면, 두 개의 부공간 사이의 관계는 식 (3.24)와 같다.

$$\begin{aligned} & V_1 V'_0 \oplus W_0 \\ & V_2 V'_0 \oplus W_0 \oplus W_1 \\ & \quad \vdots \quad \quad \quad \vdots \end{aligned} \tag{3.24}$$

이를 확장하면

$$L^2 V'_0 \oplus W_0 \oplus W_1 \oplus W_2 \cdots \tag{3.25}$$

이것은  $L^2(\mathbb{R})$  공간내의 어떤 함수  $f(x)$ 가 웨이블릿 함수의 선형 조합으로 나타낼수 있다는 것을 의미한다.

## 제 4장 제안된 시스템의 구현.

본 장에서는 3장에서 설명한 방법을 이용하여 새로운 얼굴 검출과 인식 시스템을 구현한다. 얼굴의 기본적인 특성인 켈라 정보를 가지고 얼굴 후보 영역을 획득하는 방법이 많이 적용되어지고 있다. 하지만 켈라 정보는 조명의 영향과 피부 색과 유사한 색상이 배경에 존재 한다면 얼굴 영역획득에 어려움이 존재한다. 그래서 본 논문에서는 조명과 기타 환경에 무관한 그레이 레벨에서의 얼굴 검출및 인식을 하는 알고리즘을 그림 4-1과 같이 제시한다.

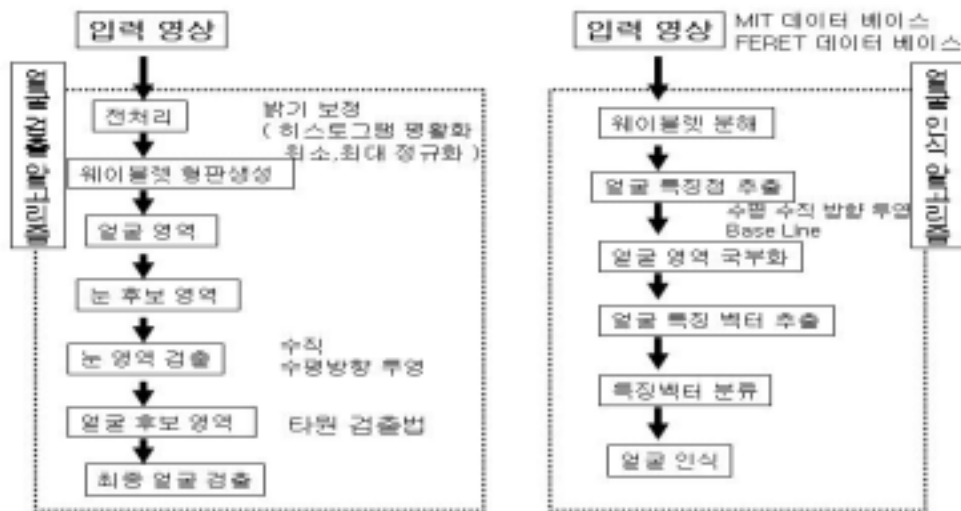


그림 4-1 얼굴 검출 및 인식 알고리즘

우선 입력 영상으로부터 얼굴 검출의 첫 단계에서 조명에 의해 얼굴 검출이 곤란한 경우들에 대해 전처리부에서 처리한다. 웨이블릿 형판을 이용하여 얼굴 영역을 설정하고, 수직 수평방향 투영을 통해 좌우 눈을 검출한다. 눈의 위치 정보를 가지고 타원형 정보로 최종 얼굴을 검출하며,



얼굴 검출과 별도로 얼굴 인식 알고리즘은, 데이터 베이스 영상을 이용하여 웨이블릿 패킷 2단계 변환을 통해 얻어진 각 16개의 얼굴영상에서 저주파영상은 3개의 평균과 분산값, 그리고 나머지 15개의 구체적 방향 영상에서 평균, 분산값으로 이루어진 총 18개의 특징벡터를 통해 유사도를 측정하여 얼굴 인식 과정을 수행한다.

#### 4.1 얼굴 검출 알고리즘

얼굴 검출 알고리즘은 그림 4-2에서와 같이 요약 할수 있다. 우선 조명에 대한 영향을 최소화 하기 위해서 앞장에서 설명한 최대-최소 정규화 과정과 히스토그램 평활화를 통한 전처리 과정을 수행한다. 그 후 수평 수직 방향 투영을 통해 눈의 영역을 검출한후 그 위치 정보를 통해 얼굴의 타원형적인 검출방법으로 최종적인 얼굴 검출 과정이 이루어 진다.

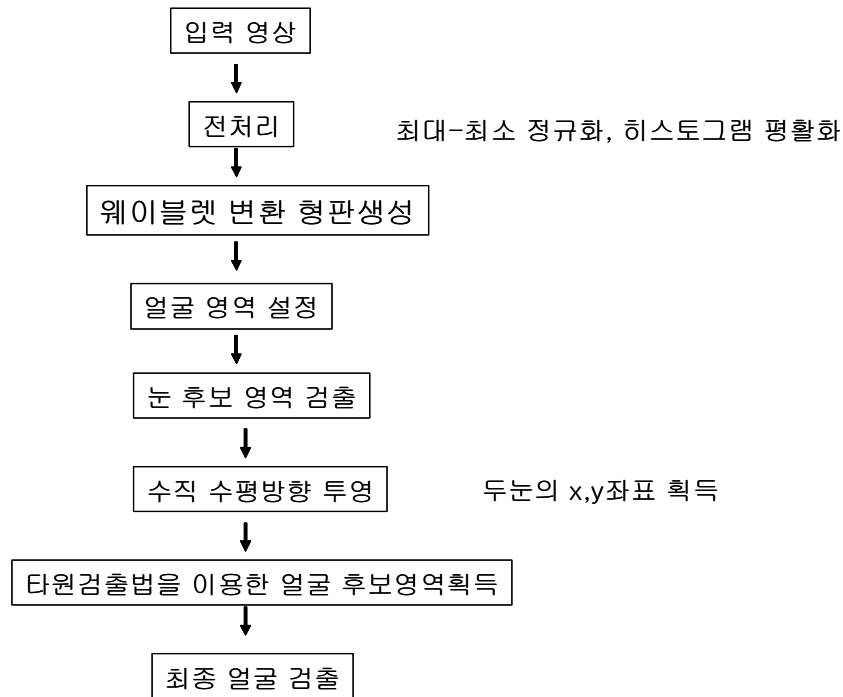


그림 4-2 얼굴 검출 알고리즘.

#### 4.1.1 전처리

입력된 영상은 256 x 256 Raw File로 그레이 영상이다. 제안된 영상의 밝기가 너무 흐리거나 너무 어두운 경우에는 얼굴 영역의 오검출이나 놓침 현상이 발생된다. 그래서 전처리 과정으로 최소-최대 정규화와 히스토그램 평활화를 수행하여 밝기의 보정을 실시한다.

입력 영상은 조명에 따라 여러 가지 형태의 입력 영상이 들어오지만 일반적인 형태로 다음 표4-1과 같이 세가지로 형태만 살펴 보기로 한다.

영상에 밝기 보정값을 위해서 최소-최대 정규화 방법을 사용하여 선형적인 변형을 취한다.

$$I' = \left( \frac{I - B_{\min_1}}{B_{\max_1} - B_{\min_1}} \right) (B_{\max_2} - B_{\min_2}) + B_{\min_2} \quad (4.1)$$

위 식에서  $B_{\min_1}$ ,  $B_{\max_1}$ 은 입력영상의 최소, 최대 밝기 값을 나타낸다.

$B_{\min_2}$ ,  $B_{\max_2}$ 는 새로운 범위의 최소, 최대 밝기 값이다.

본 연구에서 0~255값의 분포로 볼 때 상위 약 12%와 하위 약 12%값을 제외한 범위 즉 31~224 ( $B_{\min_2}=31$ ,  $B_{\max_2}=224$ )의 새로운 범위의 밝기값을 가지게 된다.

조명 상태	영상 상태
적당한 조명	일정한 조명으로 인한 검출 및 인식이 용이한 상태
너무 밝은 조명	밝기값이 너무 높아 전체적인 윤곽선 검출이 곤란한 상태
너무 어두운 조명	밝기값이 너무 낮아 전체적인 윤곽선 검출이 곤란한 상태

표 4-1 조명에 따른 영상 상태

그러나 밝기값이 180이상인 너무 높은 경우에는 단순히 위의 과정만을 가지고 전처리를 하는 경우 얼굴의 형태를 표현하기에는 부족하기 때문에 히스토그램 평활화 과정을 동시에 수행한다.

전체적인 밝기값이 180이상으로 높은 경우에는 히스토그램 평활화를 적용한 후 영상의 밝고 어두움을 강조한 후 최소-최대 정규화 과정을 적용하게 되면 영상의 밝기값이 조절되어 얼굴 검출의 오검출 현상 및 놓침 현상을 방지할수 있다.

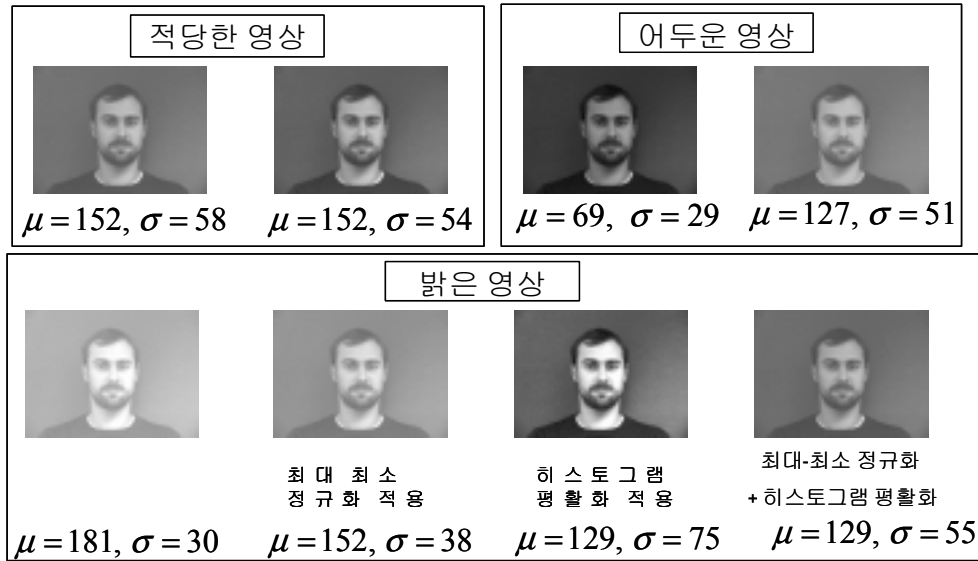


그림 4-3 제안된 밝기 조정후의 영상들

#### 4.1.2 얼굴 영상의 웨이블릿 변환

웨이블릿 변환은 사람이 사물을 바라볼 때 우선 전체적인 윤곽을 파악하고 차츰 자세한 부분을 보기 위해 집중한다는 인간의 시각 체계를 반영하는 방법으로, 서로 다른 해상도와 주파수를 적용해 영상을 처리하는 기법이다. 영상의 공간 및 주파수 영역을 동시에 표현 방법으로 압축 ( JPEG2000 ) 및 지표면 분석등 넓은 분야에서 적용되고 있다.

특히 빠른 계산 속도로 게이트 구성이 쉬운 덧셈과 곱셈으로 이루어져 있어 하드웨어 구성도 용이하다. 본 논문에서는 Daubechies 웨이블릿 필터를 사용하여 영상을 2단계로 분해한다. 얼굴 검출시 저해상도 영상을 이용하게 되면 얼굴 검출을 위한 계산량과 시간을 줄일수 있는 장점이 있으나 해상도가 너무 낮게 되면 영상의 정보가 많이 손실되게 되므로 정확한 얼굴 검출이 불가능하게 된다.

웨이블렛 변환을 얼굴 검출에 응용한 사례는 웨이블렛 변환이 얼굴 검증에 사용된 연구들을 볼수 있다 [ 53,54,55 ]. 모두 색상정보를 이용하여 얼굴 후보 영역을 검출한후, 웨이블렛 변환된 부영역을 분석하여 최종 얼굴 영역을 검증하는 방법을 사용하였다.

Karlekar, J[53]은 얼굴 후보 영역의 고주파 성분을 조사하여 충분한 고주파 성분이 존재하지 않는 후보는 무시하도록 하였으며, Christophe Garcia[55]는 웨이블렛 변환된 저주파 영역의 밝기 분산을 하나의 특징으로 삼아 얼굴 영역과 비얼굴 영역을 분류하도록 하였다. 그리고 Yanjiang Wang[54]는 웨이블렛 변환된 얼굴 영역의 수평 고주파 성분을 이진화 하여 얼굴 모형들을 생성해 놓고 이모형과 유사한 얼굴 후보 영역을 최종 얼굴 영역으로 검출하고 있다. 그러나 세 연구 모두 1차로 색상 정보를 이용하기 때문에 심한 빛의 변화나 조명의 영향과 같은 스펙트럼화된 색상의 변화등 다양한 조명 조건을 극복하지 못하였다. 또한Yuen.P.C[56]에서는 얼굴 인식에 웨이블렛 변환을 이용한 연구를 볼수 있는데, 기본적으로 PCA를 이용하여 얼굴을 인식하는 연구이지만, 입력 영상 자체의 얼굴 패턴에 대한 주성분 분석을 수행하는 것이 아니라 웨이블렛 변환된 대각선 방향 고주파 영역에 주성분 분석을 수행하여 얼굴을 인식한다.

일반적으로 얼굴 형판은 인간의 얼굴들 중 가장 일반적인 특징들을 모두 포함해야 하고 또한 동시에 배경이나 개개인의 머리모양이나 수염, 안경 착용등의 특징들에 무관하게 설정해야 한다[9,10]. 이러한 사실에 입각해서 MIT 데이터 베이스로부터 200명에 각 3장씩 총 600개의 얼굴을 형판생성에 사용하였다. 입영역은 입의 모양이 변화하기 때문에 입의 영역을 제외한 눈썹에서부터 입술위부분까지 포함하였다.

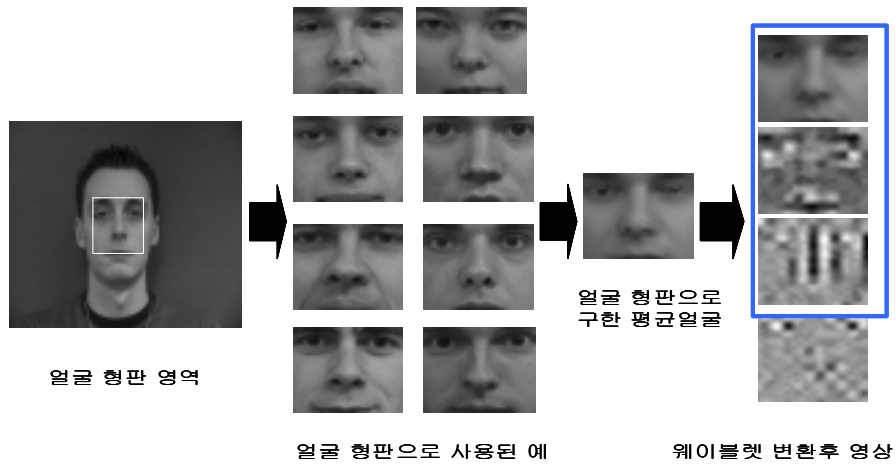


그림 4-4 얼굴 형판 생성과정

전체 영상  $256 \times 256$  중 얼굴 영상이 포함된 부영상의 크기는  $40 \times 40$ ,  $60 \times 60$ ,  $80 \times 80$ 로 웨이블릿 1단계를 적용하면 각각  $10 \times 10$ ,  $15 \times 15$ ,  $20 \times 20$ 로 축소된 영상을 얻게 된다.

웨이블릿 1단계 변환하면 저주파, 수평, 수직, 대각선 방향성분을 가진 영상으로 나타내어진다. 그러나 눈과 코 그리고 입과 같은 특징은 저주파, 수평, 수직 영상에서 특징을 검출할수 있기 때문에 대각선을 제외한 저주파, 수평, 수직 방향 성분을 가진 세가지 영상만 사용하게 된다. 즉 각 크기( $10 \times 10$ ,  $15 \times 15$ ,  $20 \times 20$ )에 따라 저주파, 수평, 수직 영상들로 총 9개의 형판을 구성한다.

#### 4.1.3 형판정합을 이용한 얼굴 검출

입력 영상에 2단계 웨이블릿 변환을 적용하게 되면 부영역에서의 1개의 픽셀이 원영상에  $4 \times 4$  픽셀에 대응되게 되므로 실제  $40 \times 40$  크기의 얼굴 영역은 웨이블릿 변환후  $10 \times 10$ 으로 작아진다. 입력 영상으로부터 다해상도 얼굴 형판을 이용하여 얼굴 검출하는 방법은 다음과 같다.

- ① 조명보정된 그레이 Raw 영상에 웨이블릿 변환하여 저주파 영역, 수평방향 고주파 영역, 수직방향 고주파 영역을 확보한다.
- ② 생성된 부영역과 형판의 유사도를 측정한다. 형판 크기(40×40, 60×60, 80×80)를 고려하면 한 프레임당 3번의 과정을 수행한다. 각 크기에 따라 3개의 형판이 존재하나 한 정합시간에 이들 세종류의 형판이 모두 정합하기 때문에 실제 정합의 횟수는 세 번이 된다.
- ③ 유사도가 가장 작은 얼굴을 검출한다.

유사도 측정방법은 Mean Absolute Error방법을 적용하였는데, 영상전체 영역을 대상으로 얼굴 후보 영역에 대해 형판을 정합한다. 그후 얼굴 영역의 중심과

형판의 중심을 맞춘 상태에서 형판의 크기가  $m \times n$ 이라 할 때,  $(m+6) \times (n+6)$ 의 탐색 영역을 설정한다. 이 안에서 얼굴 후보 영역에 대해서만 형판을 정합한다. 그리고 마지막으로 이전에 얼굴 좌표를 이용하여 얼굴의 위치를 추정하고 그 위치를 중심으로 탐색 영역을 정한다.

#### 4.1.4 눈 영역 검출 알고리즘

위와 같은 알고리즘을 수행되어진 얼굴 영역 영상에서 웨이블릿 변환된 영상 중 수평 방향 성분을 가진 영상을 가지고 수평 방향 히스토그램 으로 투영되어지면 세가지의 기준선을 얻을수 있다. 그중 상단부의 첫 번째를 눈의 y좌표로 파악하고 동일한 방법으로서 수직방향으로 투영시켜 눈의 x좌표를 검출한다. 눈의 좌우 위치가 파악되면 L.G. Frakas[52]가 제안한 식을 이용하여 얼굴 검출을 수행한다.

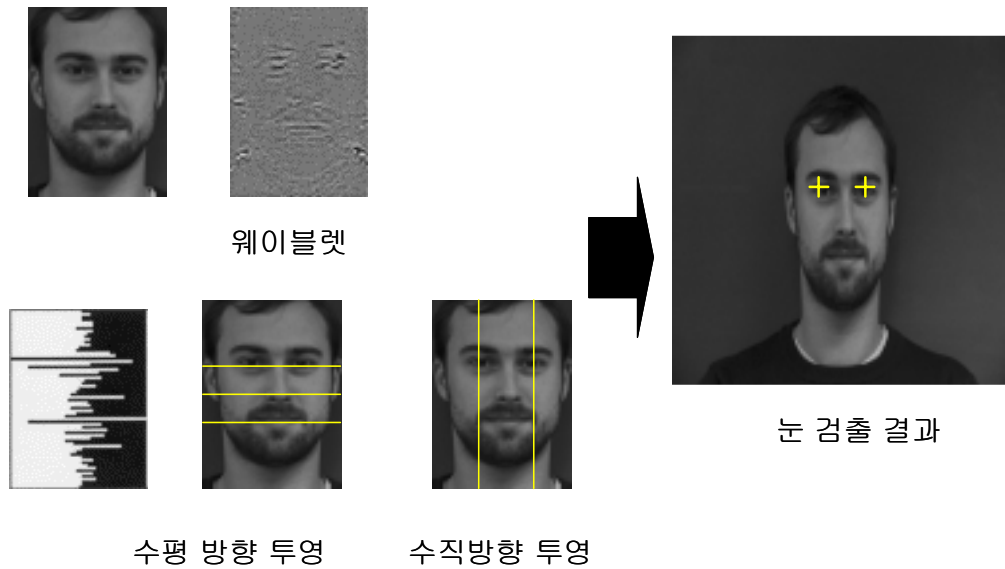


그림 4-6 눈 영역 추출

#### 4.1.5 얼굴 영역 검출

타원형 방정식으로 얼굴 영역을 설정한다. 저차원 기반에서 얼굴 검출 방법으로 모서리 검출방법은 오래전부터 사용되어져 왔다. 그러나 단순한 방법인 반면 많은 제약점을 가지고 있다. 좀더 간단한 방법으로서 웨이블릿 변환된 방향성분을 가진 영상을 이용해서 눈의 정보만을 파악하게 되면 타원형 정보를 이용하여 얼굴 영역을 검출한다.

L.G. Frakas [52]가 제안한 식을 사용하여 일반적인 사람의 얼굴 형태인 비율을 적용하게 되면 얼굴의 형태를 검출한다.

$$\frac{Height}{Width} = \frac{1+\sqrt{5}}{2} \quad (4.2)$$



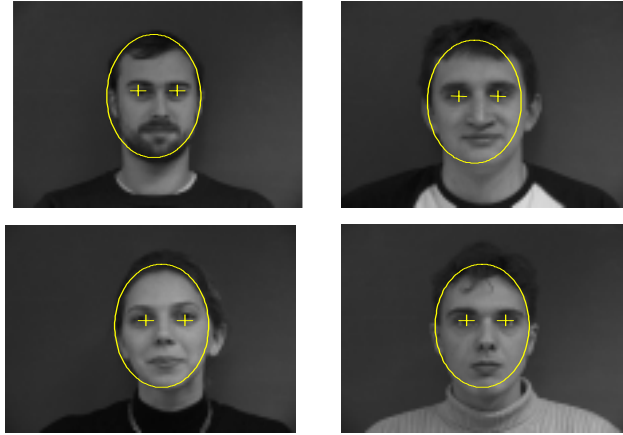


그림 4-7 웨이블릿과 Frakas식을 적용한후 얼굴 영역의 검출

## 4.2 얼굴 인식 알고리즘

얼굴 벡터를 생성하기 위해 웨이블릿 2단계 패킷 분해를 통해 실시한다. 웨이블릿 패킷 2단계를 실시하게 되면 총 16가지 영상을 생성하게 된다. 구체적인 방향 성분을 가진 15개의 영상에서 평균값과 분산값을 가진 특징 벡터와 나머지 한 개인 저주파 영상을 3가지 영역으로 분리한 평균값과 분산값을 가진 3개의 특징 벡터 총 18개의 특징벡터를 형성한다.

빠른 계산을 통해 특징 벡터들을 구분화 지을수 있는 유클리디언 거리 (Euclidian distance)를 이용하여 인식과정을 수행 한다. .

### 4.2.1 웨이블릿 패킷 분해( Wavelet Packet Decomposition )

얼굴 영상의 분석을 위해 웨이블릿 패킷 분해(Wavelet Packet Decomposition)을 수행한다. 웨이블릿 패킷 분해는 계수 행렬의 형태로서 전장에서 설명한바와 같이 시간과 공간적인 분해를 동시에 처리함으로서 영상의 다중 분석이 가능하다.

## 4.2.2 얼굴영상에서의 웨이블릿 패킷분해

1차원 신호의 웨이블릿 변환은 다음식 (4.4)과 같이 정의 된다.

$$(W_a f)(b) = \int f(x) \psi_{a,b}(x) dx, \quad \psi_{a,b}(x) = \frac{1}{\sqrt{a}} \psi\left(\frac{x-b}{a}\right) \quad (4.3)$$

$\psi$ 는 모 웨이블릿으로 영점-평균 (Zero-mean) 함수로 나타내어져야 한다. 이 제 약요소는 원영상의 다해상도 표현을 위해 모 웨이블릿  $\psi$ 에 반영된다.

2차원으로 확장한 경우에는 영상의 분리된 필터 बैं크를 적용하는데, 고주파 대역 통과 필터  $H$ 와 저주파 대역 통과 필터  $G$ 가 사용된다. 고주파 대역 통과 필터로서의 컨볼루션은 근사화 영상을 나타내고, 저주파 대역 통과 필터의 컨볼루션은 특별한 방향에 구체적인 영상을 나타낸다. 이 웨이블릿 분해에서 영상은 각각 근사화된 영상과 구체적 방향 성분을 가진 영상으로 나누어진다. 분해 과정에서 먼저 입력 영상을 수평방향에 대해서 각각 저주파 대역 통과 필터링과 고주파 대역 통과 필터링한 후 저주파 대역과 고주파 대역으로 분해한다.  $n$ 차원 분해는 다음과 같은 방법으로 나누어진다.

$$\begin{aligned} A_n &= [H_x * [H_y * A_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2} \\ D_{n1} &= [H_x * [G_y * A_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2} \\ D_{n2} &= [G_x * [H_y * A_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2} \\ D_{n3} &= [G_x * [G_y * A_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2} \end{aligned} \quad (4.4)$$

이때  $*$ 는 컨볼루션 연산자,  $\downarrow 2,1$  ( $\downarrow 1,2$ )는 열(행)에 따른 부 샘플링을 나타낸다.

$A_0 = I(x,y)$ 는 원영상을 표시하는데,  $A_n$ 은 수평방향, 수직방향 고주파 대역 통과 필터에 의해 얻어지고,  $n$  크기의 근사화 영상이다.

구체적 방향 영상  $D_{mi}$ 는 구체적인 방향 (  $i=1,2,3$  수직, 수평, 대각선 )으로  $n$  크기의 구체적인 방향정보를 나타내는 영상이다.  
 원 영상  $I$ 는 몇 개의 크기를 가진 부영상으로 다음과 같이 재 표현 할 수 있다.

$$I = \{A_n, D_{mi}\} \quad (4.5)$$

본 논문에서 제안하여 구성되었던 웨이블릿 패킷 분해는 전형적인 웨이블릿 변환을 응용하여 표현하며, 근사화된 영상 뿐만아니라 구체적인 방향 영상으로 나누어 표현되어진다. 웨이블릿 분해 가지에서의 결과에 적용될  $H, G$  필터는 그림 4-8에서 같이 나타내었다.

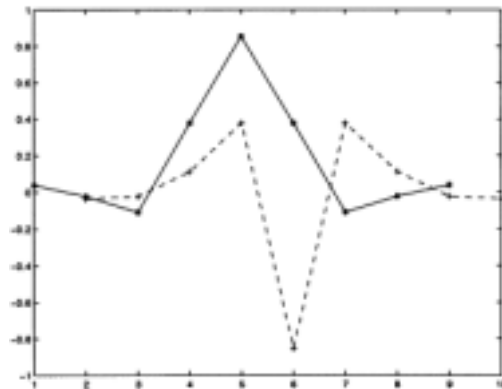


그림 4-8 웨이블릿에 적용할 H,G 필터 H(실선) G(점선)

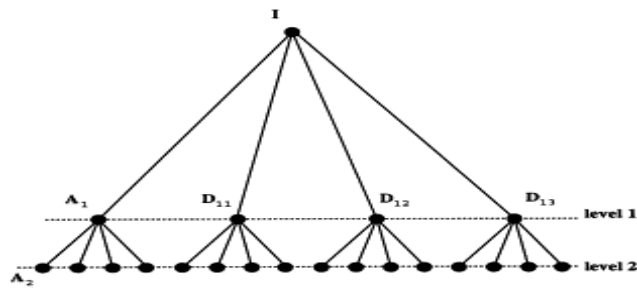


그림 4-9. 웨이블릿 패킷 가지

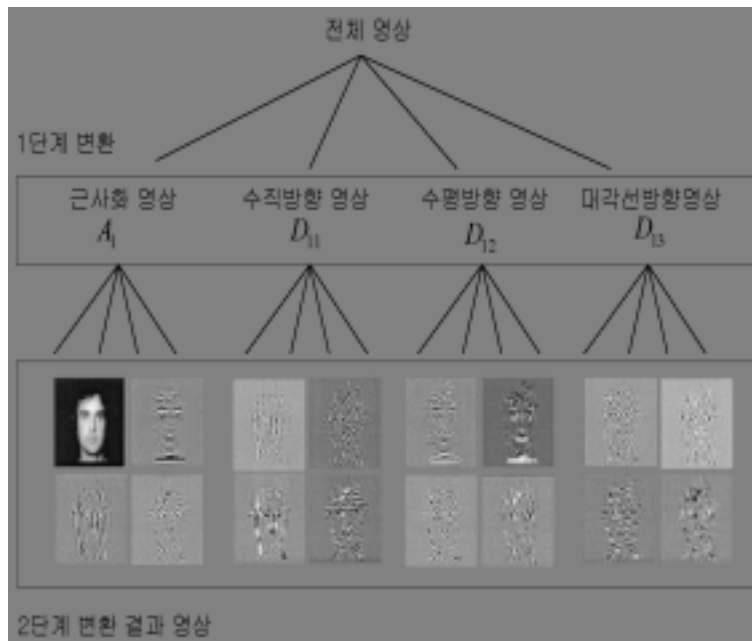


그림 4-10. 웨이블릿 패킷 가지의 2단계를 통해 얻어진 부영상들

제안된 과정은 그림 4-9에서 보여준 바와 같이 두 단계의 웨이블릿 패킷 분해를 통해 구성된다. 2단계 이상의 분해를 하게 되는 경우 영상의 크기가 작게 되어 처리 속도가 좋아질 수는 있지만 표현할 정보가 불충분하기 때문에 2단계의 분할 과정만 수행하였다. 2단계 분해로서 얻어진 영상은 그림 4-10에서와 같이 한개의 근사화된 영상과 15개의 방향성분을 가진 영상으로 표현된다.

전체 16개의 웨이블릿 계수 행렬로 표현된 얼굴 영상의 크기는 입력영상과 동일한 크기이므로 상당히 큰 양의 정보를 표현한다. 패턴 분류과정은 패턴 공간의 차수에 의해 영향을 받게 되므로 추출된 특징에 의해 판별이 가능한 필수적인 요소를 결정함으로써 차원의 감소가 필요하다. 패킷 분할된 16개의 영상은 각각 얼굴의 특징 정보를 포함한다.

#### 4.2.3 얼굴 특징 국부화.

웨이블릿 패킷 분석과 특징 추출 과정 전에 얼굴의 구체적인 영역을 설정하여 배경을 제외하도록 한다. 웨이블릿 계수로서 계산하고, 얼굴의 특징의 통계적인 정보를 통해 얼굴 영역을 추출한다.

보통 얼굴의 일반적인 특징을 살펴보면 얼굴 영역은 코를 기준선으로 하는 윗 부분과 아랫부분으로 두부분으로 나누어 진다. 그리고 얼굴의 특징에서 좌우 대칭적인 정보를 가지고 왼쪽과 오른쪽 경계를 찾는다. 두 눈의 중심과, 코, 입의 주축을 중심으로 하여 각 얼굴 특징은 대칭적이라 볼 수 있다. 얼굴의 대칭적인 면에서 정면 얼굴을 볼 때 두 눈은 수평적이다. 이런 일반적인 특징을 가지고 정면 얼굴의 특징을 추출하기 위해 많은 알고리즘들이 제시되어 졌는데, 주요 기반은 템플릿 매칭 방법이다. 템플릿은 눈과 입 그리고 코와 같은 특징 부분에 적용되어 지는데, 크기의 다양성을 해결하기 위해 각각 다른 크기의 템플릿이 요구되어 지기도 하고, 얼굴 영상내에 템플릿들을 위치 시키는데, 계산상 복잡한 기술이 요구된다. 빠른 계산적인 요구를 위해 간단한 영상 처리 과정인 히스토그램을 사용해서 얼굴 특징들의 대략적인 위치를 찾는다. 이 얼굴 특징의 추출을 위한 적분 투영(integral projection)의 기술이 Kanade에 의해 제안되어졌다. 영상  $I(x,y)$ 에서 수평,수직 투영은 각각  $H(y)$ ,  $V(x)$ 로 나타낸다.

$$\forall \{x, y\} \in A, H(y) = \sum_{x=x_1}^{x_2} I(x, y), V(x) = \sum_{y=y_1}^{y_2} I(x, y)$$

$$A = [x_1, y_1] \times [x_2, y_2] \quad (4.6)$$

적분 투영의 결과로서 얼굴 특징점의 위치를 검출한다.

Kanade의 방법에서의 적분 투영은 라플라시안 연산자와 임계값을 적용한 후 얻어진 이진화 영상에 사용되어졌지만, 제안된 연구방법은 웨이블렛 패킷분해를 통해 얻어질수 있는 대역통과 필터링된 영상들이 가지고 있는 장점을 이용하였다.

1단계 수직방향에 구체화된 방향 성분을 가진 영상  $D_{11}$ 을 가지고 수직방향으로 투영시킨다. 식(4.7)에서  $V(x)$  벡터를 투영의 결과로부터 두가지 최대값을 통해 얼굴의 왼쪽과 오른쪽의 외곽선을 찾는다. 이 왼쪽, 오른쪽 두 위치는 얼굴을 포함하는 수직 방향의 대역으로 정의한다. 이 외곽안에서 구체적인 얼굴의 특징점을 찾는다. 그리고 1단계 수평방향에 구체적 방향 성분을 가진 영상  $D_{12}$ 에서 눈과 코, 그리고 입의 위치는 수평방향으로 투영시 높은 값을 나타내는 얼굴 특징에 대한 정보를 사용하여  $H(y)$  벡터의 국부적인 최대값을 찾는다. 수평방향으로 투영하면 세가지 기준선을 얻을수 있는데, 이것은 눈, 코, 입의 위치를 나타내는 기준선이다.

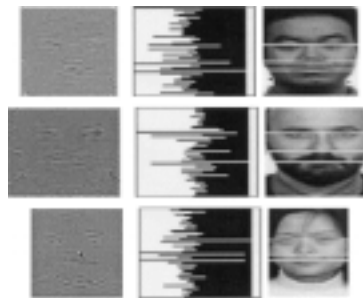


그림 4-11 FERET 데이터로부터 추출된 얼굴들

그림 4-11은 FERET 데이터로부터 수평 투영 방법으로 얼굴의 기준선을 얻은 몇가지 예를 보여주고 있다. 각각의 얼굴에서 왼쪽 영상은 웨이블릿 계수에 의한 1단계 수평 방향 성분을 가진 영상이고, 가운데 영상은 수평방향으로 투영 시켜 얻은 벡터이고, 그리고 마지막 영상은 세가지 최대값으로 선택된 기준선을 잡은후 원 영상에 표시한 그림이다.

그 후 얼굴의 외곽은 폭으로서 외곽 사각형의 형태를 결정하는데, 머리의 형태와 수염 그리고 기타 여건에 의해 얼굴의 위, 아래를 정하는 기준선을 결정하기는 곤란하다. 그러나 이런 문제점은 사람 얼굴 모양이 비교적 일정하다는 것에 기인하여 외곽 사각형의 형태를 선택하여 해결 한다.

그림 4-12에서 보여주는 바와 같이 얼굴의 폭은 수직방향 성분을 가진 웨이블릿 변환된 영역을 이용하여 결정한다. 얼굴의 외곽은 수직 투영된 결과 잘 나타난다. 그리고 수평적인 위치는 일반적인 얼굴 모양의 정보를 이용하는데, 그 특징은 눈과 코의 간격은 코와 입의 간격에 비해 약 1.5배 이고, 코를 기준선으로 잡아 눈과 코의 간격 두배의 거리를 얼굴의 상단부로 파악한다. 마찬가지로 코의 기준선에서 얼굴 영역의 상단부로 잡은 거리와 동일하게 얼굴의 하단부 영역을 설정한다. 수직 수평 영역이 결정하고 얼굴의 영역을 좀더 명확히 결정하기 위해 15%의 여유를 두어 최종 얼굴의 사각형 영역을 결정한다



그림 4-12. 저주파 얼굴의 분리후 영상

얼굴 인식에 앞서 얼굴 부분의 검출은 중요한 과정이다. 제안된 방법은 기준선의 위치에서 매우 정확한 검출이 필요하지 않기 때문에 1단계 웨이블릿 필터링된 구체적인 방향 영상을 사용하여 수행한 과정은 얼굴 영역 검출에 효과적이었다. 이 얼굴 특징 국부화 과정에서 정확한 얼굴 특징을 찾는 것을 필요로 하지 않을 뿐만 아니라 두 영역으로 얼굴을 대략 나눌 수 있다. 더욱이 좀더 낮은 해상도인 2단계에서도 특징추출이 얻어질 수 있다.

#### 4.2.4 특징 벡터 추출

얼굴 벡터는 2단계 근사화된 저주파 성분 영상에서 얼굴 내부의 상단부, 하단부 그리고 외곽 영역 세가지 영역과 방향성분을 가진 영상으로부터 특징 벡터를 추출한다. 웨이블릿 필터링된 영상으로부터 얻어진 영상은 각각의 얼굴 정보를 가지고 있다. 이 얼굴 정보는 수평, 수직방향에 따라 웨이블릿 계수에 분류에 의해 특징을 가지고 있다. 평균값과 분산값을 가지고 벡터를 통해 분류하도록 한다.

얼굴의 영역은 머리 모양이나 기타 얼굴 형태에 관한 일반적인 정보를 제공한다. 그리고 내부 영역은 상단부에 속한 눈과 눈썹정보를 그리고 하단부에는 입과 같은 정보와 피부의 명암을 제공하는데, 코 기준선으로 분리된 얼굴의 상단부분과 하단부분이다. 얼굴의 영역, 상단부분 그리고 하단부분 영역으로부터 각각 평균값 ( $\mu_{out}, \mu_{top}, \mu_{bottom}$ )과 분산값 ( $\sigma_{out}^2, \sigma_{top}^2, \sigma_{bottom}^2$ )을 추출한다.

다른 15개의 구체적인 방향 영상에서 평균값 ( $\mu_i$ )과 분산은 ( $\sigma_i^2$ ) ( $i=3, \dots, 17$ ) 전체 영상에서 추출된다. 즉 특징 벡터는 총 18개의 나타낸다.(근사화된 영상에서 3개의 평균값과 3개의 분산값 그리고 구체적 방향영상에서 15개의 분산값) 즉 다음과 같이 표현한다.

$$v = \bigcup_{i=0}^{17} \{\mu_i, \sigma_i^2\} \quad (4.7)$$



여기서  $\forall \geq 3, \mu_i = 0, (i=0,1,2$  얼굴 영역의 상단부 하단부, 외곽영역의 평균값과 분산) 이들 18개의 벡터값들을 가지고 유클리디언 거리를 이용하여 각 얼굴의 특징 벡터를 분리 하였다.

## 5장 실험 및 결과.

### 5.1 실험 환경 및 데이터 베이스

지금까지 제안한 알고리즘을 실험하기 위해서 본 논문에서는 P.J. Phillips[57]가 FERET 실험과정으로 제안한 과정으로 실험 환경은 Pentium III 128MB에서 수행하였다. 이 평가 과정은 두 가지 실험 방법들로 구성되었는데 각각 훈련에 적용되는 집합을 갤러리(gallery), 그리고 훈련 과정을 거치지 않은 프로브(probe) 집합으로 구분한다.

첫 번째 테스트는 사람 당 하나의 이미지를 사용하여 갤러리로부터 개인의 신원을 확인하기 위한 성능(Identification Performance)을 유클리디언 거리 측정 방법을 사용하여 실행한다. 두 번째 테스트는 훈련되지 않은 얼굴의 표정에 따른 얼굴 변화에 따른 인식률 테스트이다. 그리고 표 5-1은 실험 데이터와 실험 환경을 나타내었다.

학습영상수	MIT 데이터베이스	200명 총 600장
	FERET 데이터베이스	155명 총 310장
입력 영상 해상도	256×256 8bit gray 영상	
2차원 웨이블릿 패킷 변환후 해상도	16×16×16	
얼굴 영역	MIT 데이터베이스	35×50
	FERET 데이터베이스	비정규 40×40 ~ 55×55
특징벡터수	18	

표 5-1 실험 데이터와 실험 환경

본 논문에서의 실험은 두 가지 형태의 데이터 베이스를 사용하여 구현하였다. 첫 번째가 200명의 각 개인당 3가지씩 총 600개의 영상을 담고 있는 MIT의 FACES 얼굴 데이터베이스이고, 두 번째로 155명의 각 개인당 2가지씩 총 310장의 영상을 담고 있는 FERET 데이터베이스이다. FERET 데이터 베이스는 동일한 사람인 경우 얼굴 표현 다르게 하여 구성하며 잘 정규화 되지 않은 상태이다. 그림 5-1과 그림 5-2는 두가지 데이터 베이스에서의 몇몇 영상들을 보여주고 있다.



그림 5-1 MIT FACES DATA의 예



그림 5-2 FERET DATA의 예

## 5.2 개인 얼굴 식별 실험

제안된 알고리즘은 우선 전체 데이터에서 점차적으로 새로운 얼굴 영상을 증가 시키면서 실험하는데, MIT 데이터의 경우 총 600장 그리고 FERET 데이터의 경우 총 310장의 데이터를 가지고 실험했다. 각 개인의 얼굴을 우선적으로 데이터의 양을 MIT데이터는 60개에서부터 30개씩 데이터 양을 늘려가면서 그리고 FERET데이터의 경우 60개에서부터 20개씩 데이터의 양을 증가 시키면서 인식율을 비교하였는데, 특징 벡터는 각 클래스  $c$ 의 평균벡터 ( $v_c^{mean}$ )를 계산함으로써 모든 영상에서 구한다. 그리고 인식할 영상  $k$ 와 각 클래스  $c$ 의 평균벡터 ( $v_c^{mean}$ )와의 최소 거리  $D(v_k, v_c^{mean})$ 를 유클리디언 거리 측정에 적당한 클래스로 구분하여 영상  $k$ 의 특징 벡터와 각 클래스의 평균벡터의 거리  $D(v_k, v_c^{mean})$ 가 최소가 되면 클래스  $c$  ( $v_c^{mean}$ )에 속하게 된다.

표 5-2와 표 5-3은 결과를 보여 준다.

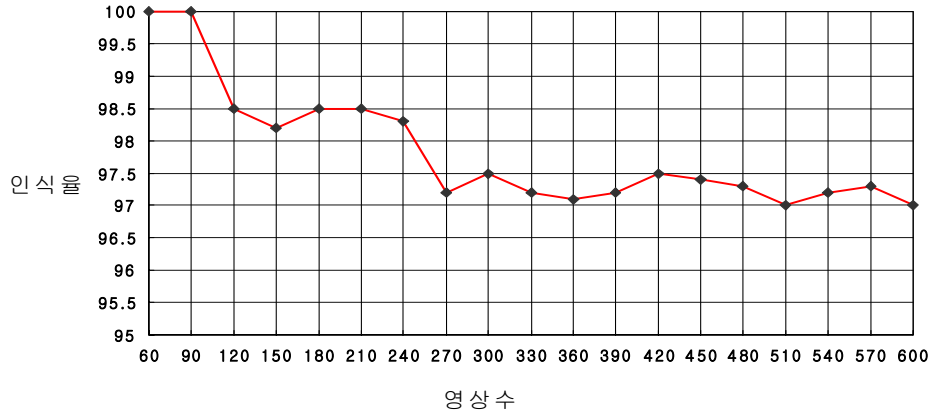


표 5-2. FACES 영상 수에 따른 인식률

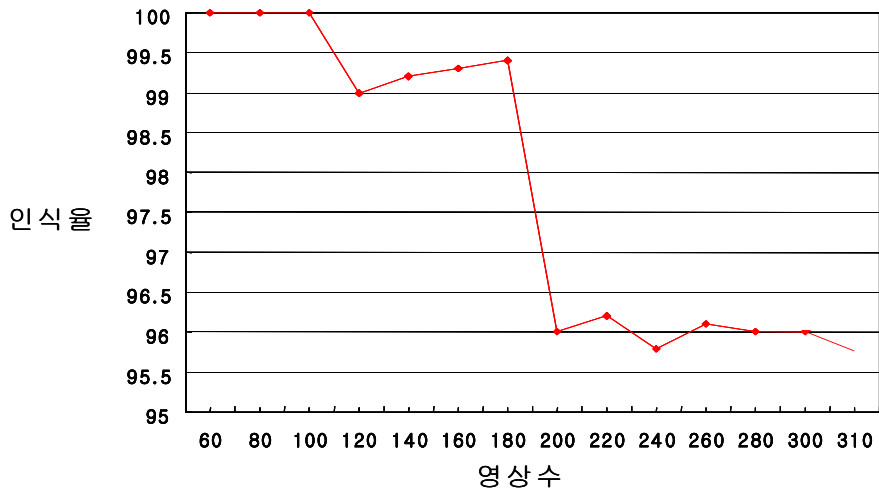


표 5-3. FERET 영상에 따른 인식률

FACES 영상은 100%(90/90) ~ 97%(582/600)의 인식 성공률을 보였고, FERET 영상은 100%(100/100) ~ 95.8%(297/310)의 인식 성공률을 보였다. FACES 데이터 집합에서 완전한 100%의 인식율은 90개의 영상까지로 FERET데이터는 100개 까지 실행하였다.

### 5.3 얼굴 표정 변화시 인식을 실험

두 번째 실험은 MIT FACES 데이터 영상만을 이용하여 실험하였다. FERET 데이터 영상의 경우 각 개인당 2장의 영상만을 가지고 있으므로, 훈련 데이터를 사용하기에는 무리가 있다고 판단된다. 그래서 각 개인당 3장의 영상을 가지고 있는 MIT 데이터 베이스를 이용하였는데 그림 5-1에서 보여준 것과 같이 처음 두가지 영상은 훈련 영상으로서 평균벡터를 계산하기 위해 훈련 데이터로서 처리 하였고, 마지막 세 번째 영상은 테스트 영상으로서 처리하였다. 결과는 표 5-4에서 나타내었는데, 120개의 영상수에서 현저히 낮은 인식을 보이고 있고, 전체 데이터의 인식률은 90.7%(281/310)까지 감소하였다.

이것은 각각의 클래스에서 얼굴의 다양하게 변화함에 따라 단지 두가지 영상만을 사용하는 것은 좋은 평균벡터를 생성하기 곤란하고, 다양한 훈련 데이터를 통해 평균 클래스 벡터를 정하여야 한다는 것을 의미하기 때문에 차후에 많은 훈련데이터의 양을 확보하여 실험하여야 해서 인식을 향상시키도록 하여야 한다..

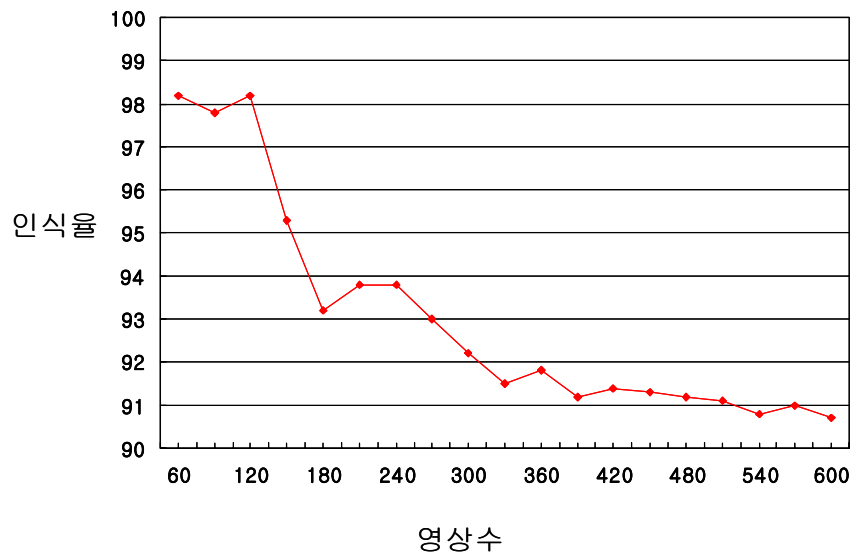


표 5-4 MIT 데이터 베이스의 얼굴 표정 변화 입력시 인식률

## 5.4 PCA 방법과의 실험 비교

고유얼굴을 이용한 연구의 경우, 각 얼굴영상은 고유얼굴들의 합인 고차원적인 특징 벡터로 다루어지고 있는데, 각 영상은 고차원 공간내에서 간단한 점들로 각 픽셀마다 하나의 특징을 가지고 있기 때문에 특징 벡터의 차원은 작은 영상의 크기인 경우에도 매우 크다.

PCA를 사용한 고유얼굴 방법은 영상 공간을 저 차원 공간으로 한 선형 변화해서 저차원 공간의 결과로 직교한 기저 벡터가 고유 얼굴이 되고 이를 저장한다. 인식하기 위한 각각의 얼굴은 이 고유 얼굴의 각각의 투영되고 결과의 특징 벡터가 각각의 요소에 따라 주어진다. 유클리디언 거리는 특징 벡터를 분류하기 위해 사용하였다.



그림 5-3 FACES와 FERET 영상의 여섯개의 고유얼굴들

그림 5-3에서 첫 번째 여섯 개는 FACES영상 데이터에서 계산된 고유 얼굴이다. 그리고 아래 여섯개는 FERET 데이터 영상에서 얻은 결과이다.

	PCA 방법	제안한 방법
MIT FACES 데이터	98.17% (589/600)	97%(582/600)
FERET 데이터	70.32% (218/310)	95.8%(297/310)

표 5-5 PCA방법과 제안한 방법과의 비교

고유얼굴 방법은 두가지 데이터에 적용하였는데, FACES 데이터 영상에서는 좋은 결과를 얻을수 있었다. 이것은 PCA방법을 위해 잘 형성된 영상들의 집합이기 때문이다. 제안된 방법에서 사용한 97% (582/600)와 비교하여 180개의 고유얼굴을 사용하여 98.17%(589/600)결과를 얻었다. FERET 데이터의 경우 영상은 FACES영상들 처럼 표준화 되어 있지 않기 때문에, 150개의 고유얼굴을 사용하여 70.32%의 성공률을 보였지만 PCA를 사용하는 경우 FERET 데이터 베이스의 얼굴 영역만을 가지고 실험 한다면 좋은 효과를 나타낼 수 있을 것이다. 하지만 고유얼굴 사용방법의 문제점은 고유얼굴을 저장하여야 하므로 데이터베이스에 여분의 공간을 두어야 하고, 두 번째 문제점으로는 FERET 데이터의 경우 처럼 입력 영상은 표준화되어야 한다.



## 제 6 장 결론 및 추후 과제

---

얼굴 인식 알고리즘에서 제안된 과정은 정규화 과정이 필요 없고 얼굴의 대략적인 기준선을 통해서 얻어진 국부화된 영상을 통해 특징 벡터를 생성 분류한다. 또한 PCA방법과 달리 고유 얼굴의 계산과 저장과정이 없다는 것이다.

제안된 과정에서 얻어진 결과는 정규화 되지 않은 FERET실험 데이터 영상의 경우 PCA방법을 적용한 경우 보다 우수했고, 고유얼굴을 비교하고, 인식하기 위해 오랜 시간이 걸린 PCA에 비해 본 논문에서 제안한 방법은 적은 시간을 요하고 있다.

그러나 PCA의 방법의 문제점인 전체적인 얼굴 영상 정보에 기반하여 얼굴을 모델링한다는 단점을 보완하기 해서 국부적인 특징을 분석 (Local Feature Analysis)라는 방법이 적용되고 있는 점을 감안할 때, 얼굴 인식에 우선 단계인 검출 단계와 인식 단계를 체계화하여 하나의 시스템으로 실용화하기 위해서는 앞으로 계속된 연구가 진행되어야 한다.

그리고 본 논문에서는 배경이 없고 얼굴 표정의 약간의 변화만 가지고 있는 두가지 데이터 베이스를 통해서 얼굴 인식 문제를 다루었는데, 웨이블릿 변환을 실시한 경우 주변의 배경의 영향을 받게 되므로 이 또한 배경이 존재시 배경을 제외할수 있는 실질적인 응용을 위한 연구가 필요하다.

그리고 5.3절에서와 같이 훈련된 데이터에 얼굴 표정이 변화된 영상을 입력시 그 영상을 인식하기 위해서 훈련되어야 하는 데이터의 수가 부족하여 얼굴 인식이 저하되는 것을 볼수 있었다. 그런 문제점들을 해결하기 위해서는 많은 데이터를 확보하여 훈련 시켜야 한다.

## 참고 문헌 ( Reference )

---

- [1] R.Chellappa, Charles L. Wilson, and Sirohey, "Human and Machine Recognition of Faces: A Survey," Proc IEEE, Vol.83, No.5 pp.704~740, May 1995
- [2] M.Turk and A. Pentland, "Eigenfaces for Recognition," Journal of Cognitive Neuroscience, pp71~86, 1991.
- [3] Shigeru Akamatsu, Tsutomu Sasaki, and Hideo Fukamachi, "An Accurate and Robust Face Identification Scheme," Proceedings of IEEE, pp.217~220, 1995.
- [4] I.Craw, H.Ellis,and J.R. Lishman, "Automatic extraction of face-feature", Pattern Recognition, Lett. Feb. 1987
- [5] T.Sakai, M.Nagao,and T.Kanade, "Computer analysis and classification of photographs of human faces", in Proc. First USA-Japan Computer Conference pp2~7, 1972
- [6] X. Jiwang, M.Binkert, B.Achermann,and H.Bunke, "Towards detection of glasses in facial images, Pattern anal". pp9~18, Appl 3, 2000,
- [7] Z.Jing and R.Mariani, "Glasses detection and extraction and extraction by deformable contour", in Processings of the 15th International Coference on Pattern Recogniton, 2000
- [8] C.Wong, D.Kortenkamp, and M.speich, "A mobile robot that recognises people", in IEEE Int. Conf. on Tools with Artificial Intelligence, 1995
- [9] R.Hoogenboom and M.Lew, "Face detection using local maxmia", in IEEE Proc. of 2nd Int. Conf. on Automatic Face and Gesture Recognition, 2000
- [10] J.Yang and A.Waibel, "A real-time face tracker", in IEEE proc. of the 3rd Workshop on Applications of Computer Vision, Florida, 1996

- [11] F.Luthon and M.Liecin, Lip "Motion automatic detection", in Scandinavian Conference on Image Analysis, Lappeenranta, Finland 1997
- [12] B.K.Low and M.K.Lbrahim, "A fast and accurate algorithm for facial feature segmentation", in Proceedings International Conference on Image Processing 1997
- [13] D.Reisfeld, H.Wolfson, and Y.Yeshurun, "Context-free attentional operators: The generalized symmetry transform", *International J. Comput. Vision* 14, pp119~130, 1995
- [14] L.C.De Silva, K.Aizawa,and M.Hatori, "Detection and tracking of facial features by using a facial feature model and deformable circular template", *IEICE Trans. Information System* E78-D(9), pp1995-1207, 1995
- [15] M.C Bual, T.K. Leung, and P.perona, "Face localization via shape statistics", in *Int. Workshop on Automatic Face and Gesture Recognition*, Zurich, Switzerland, June 1995
- [16] M.Kass, A.Witkin,and D.Terzopoulous, "Snakes: active contour models", in *Proc. of 1st Int Conf. On Computer Vision*, London, 1987
- [17] A.L.Yuille,P.W.Hallinan, and D.S. Cohen, "Feature extraction from faces using deformable templates, *Int". J. Comput. Vision.* pp99~111, 1992
- [18] A.Lanitis, C.J.Taylor, and T.F. Cootes, "Automatic tracking, coding and reconstruction of human faces, using flexible appearance models", *IEEE Electron. Lett.*30, pp1578~1579, 1994
- [19] Y.Saito, Y.Kenmochi, and K.Kotani, "Extraction of a symmetric object for eyeglass face analysis using active contour model", in *Proceeding of the 2000 International Conference on Image Processing*, 2000
- [20] L.Sirovich and M.kirby, "Low-dimensional procedure for the characterization of Human faces", *J.Opt. Soc. Amer.* pp519~524, 1987
- [21] M.Turk and A.Pentland, "Eigenfaces for recognition", *J.Cog. Neurosci.*

- pp71-86, 1991
- [22] R.A.Fisher, "The use of multiple measurements in taxonomic problems",  
Ann.Eugenics, pp179~188, 1936,
- [23] Pentland, B. Moghaddam, and T. Strarner, "View-Based and modular  
eigenspaces for face recognition, in IEEE Proc. of Int. Conf. On  
Computer Visionand Pattern Recogniton, 1994
- [24] B.Moghaddam and A.Pentland, "Probabilistic visual learning for object  
representation" IEEE Trans. Pattern Anal. Mach. Intell 19(1) 1997.
- [25] T.S Jebara and A.Pentland, "Parameterized structure from motion for 3D  
adaptive feedback tracking of faces", IEEE proc. of Int.Conf. On Computer  
Vision and Pattern Recognition. 1997
- [26] L.Meng and T,Nguyen, "Two subspace methods to discriminate faces and  
clutters", in Proceeding of the 2000 International Conference on Image  
Processing, 2000
- [27] K-K. Sung and T.Poggio, "Example-based learning for view-based human  
face detection", IEEE Trans. Pat. Anal. Mach. Intelligence pp39~51, 1998
- [28] E.Viennet and F.Fogelman Soulie, "Connectionist method for human face  
processing in Face Recognition" From Theory to Application. 1998
- [29] H.A. Rowley, S.Baluja, and N. Ahuja, "A SNoW-based face detetion",  
IEEE Trans. Pattern Anal Mach. Intell. pp23~38, 1998
- [30] R.Feraud, O. Bernier, and D.Collobert, "A fast and accurate face detector  
for indexation of face images", in Proceedings Fourth IEEE International  
Conference on Automatic Face and Gesture Recognition, 2000
- [31] S-H.Lin, S-y. Yang, and L-J Lin "Face recognition/detection by  
probabilistic decision-based neural network" IEEE Trans.  
Neural Networks pp114~132, 1997
- [32] D.Roth, M-H. Yang, and N. Ahuja, "A SNoW-based face detector in  
Advances" Neural Information Processing Systems 12, MIT Press,

Cambridge, MA,2000

- [33] E.Osuna, R.Freund, and F. Girosi, "Traning support vector machines  
An application to face detection", in IEEE proc. of Int. Conf. On Computer  
Vision and Patt. Recog. 1997
- [34] V.Vapnik, "The Nature of Statistical Learning Theory" New York, 1995
- [35] M.Nixon, "Eye Spacing Measurement for Facial Recognition" SPIE Proc.  
Vol. 575, pp279~285, 1985
- [36] Roberto Brunelli and Tomaso Poggio. "Face Recognition: Feature versus  
Templates." IEEE Trans. On PAMI, Vol 15, No 10, 1993
- [37] A.Yuille, D.Cohen, and P.Hallinan, "Feature Extraction from Faces using  
Deformable Template" Proc IEEE Conf. On Computer Vision and Patt.  
Recog, pp104~109, 1989
- [38] A. Samal and P.A. Iyengar, "Automatic recognitionand analysis of human  
faces and facial expression : A survey", Pattern Recog. pp65~77, 1992
- [39] S.A Sirohey "Human Face Segmentation and Identification"  
Tech.Rep CAR-TR-695, Center for Autom Res. Univ.Maryland, 1993
- [40] D.H Ballard "Generalizing the Hough Transform to detect arbitrary shapes"  
Patt. Recog Vol 13. pp111~122, 1981
- [41] S.D. Shapiro "Feature Space Transform for Curve Detection"  
Patt Recog Vol 10, pp129~143, 1978.
- [42] J.Illingworth and Kittler ,"The Adaptive Hough Transform"  
IEEE Trans on PAMI Vol 9, NO 5, pp690~698, 1987
- [43] X,Li and N. Roeder, "Face Contour Extraction from Front-View Images"  
Patt. Recog Vol 28, No 8, pp1167~1179, 1995
- [44] P C Yuen, D Q Dai and G C Feng "Wavelet-based PCA for Human Face  
Recognition"Dep.of Computing Studies,Hong Kong Baptist Univ. IEEE  
pp223~228 1998
- [45] Feris, R.S., Cesar, R.M., Jr "Tracking facial features using Gabor wavelet

- networks" Computer Graphics and Image Processing, 2000.
- [46] Y.Zhu, S.Schwartz, and M.Orchard, "Fast Face detection using subspace discriminant wavelet features", in IEEE Conference on Computer Vision and Patt. Recognition, 2000
- [47] Feris, R.S. Cesar, R.M., Jr. Kruger, V. "Efficient real-time face tracking in wavelet subspace Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems" Proceedings. IEEE ICCV Workshop pp 113~118, 2001
- [48] C.SBurrus, R.A.Gopinath, H.Guo, "Introduction to wavelets and Wavelet Transform" A Primer, Prentics Hall,
- [49] C.K.Chui, "An Introduction to Wavelets". New York, Academic Press, 1992
- [50] I.Daubechies, "Ten Lectures on Wavelets" PA: SIMA Press, 1992.
- [51] J.Zhang, G.G.Walter, Y.Miao, and N W Lee, "Wavelet Neural Network for Function Learning" IEEE Trans.Signal processing, Vol.43, No6, pp1485~1497, June 1995
- [52] L.G. Frakas and I.R.Munro, "Anthropometric Facial Proportions in Medicine" Charles C. Thomas, 1987

# ABSTRACT

## Face Detection and Recognition Using Ellipsodal Information and Wavelet Packet Analysis

Jung, Myungho

Dept. of Biomedical Engineering

( Electrical and Electronic Eng )

The Graduate School

Yonsei University

---

This paper deals with face detection and recognition using ellipsodal information and wavelet packet analysis. We proposed two methods.

First, Face detection method uses general ellisodal information of human face contour and we find eye position on wavelet transformed face images.

A novel method for recognition of views of human faces under roughly constant illumination is presented.

Second, The proposed Face recognition scheme is based on the analysis of a wavelet packet decomposition of the face images. Each face image is first located and then, described by a subset of band filtered images containing wavelet coefficients. From these wavelet coefficients, which characterize the face texture, the Euclidian distance can be used in order to classify the face feature vectors into person classes.

Experimental results are presented using images from the FERET and the MIT FACES databases. The efficiency of the proposed approach is analyzed according to the FERET evaluation procedure and by comparing our results with those obtained using the well-known Eigenfaces method. The proposed system achieved an rate of 97%(MIT data) ,95.8%(FERET databace)

---

Key Words : Face Detection, Face Recognition, Wavelet Transform,  
Ellipsodal Information, Bioinfomation Recognition